THE JOURNAL OF FINANCE • VOL. LVIII, NO. 4 • AUGUST 2003

Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps

RAVI JAGANNATHAN and TONGSHU MA*

ABSTRACT

Green and Hollifield (1992) argue that the presence of a dominant factor would result in extreme negative weights in mean-variance efficient portfolios even in the absence of estimation errors. In that case, imposing no-short-sale constraints should hurt, whereas empirical evidence is often to the contrary. We reconcile this apparent contradiction. We explain why constraining portfolio weights to be nonnegative can reduce the risk in estimated optimal portfolios even when the constraints are wrong. Surprisingly, with no-short-sale constraints in place, the sample covariance matrix performs as well as covariance matrix estimates based on factor models, shrinkage estimators, and daily data.

MARKOWITZ'S (1952, 1959) PORTFOLIO THEORY is one of the most important theoretical developments in finance. Mean-variance efficient portfolios play an important role in this theory. Such portfolios constructed using sample moments often involve large negative weights in a number of assets. Since negative portfolio weights (short positions) are difficult to implement in practice, most investors impose the constraint that portfolio weights should be nonnegative when constructing mean-variance efficient portfolios.

Green and Hollifield (1992) argue that because a single factor dominates the covariance structure, it would be difficult to dismiss the observed extreme negative and positive weights as being entirely due to the imprecise estimation of the inputs. They consider the global minimum variance portfolio to avoid the effect of estimation error in the mean on portfolio weights. They note that when returns are generated by a single factor model, minimum variance portfolios can be constructed in two steps. First, naively diversify over the set of high beta stocks and

*Jagannathan is from the Kellogg School of Management at Northwestern University and the National Bureau of Economic Research and Ma is from the David Eccles School of Business at the University of Utah. We thank Torben Andersen, Gopal Basak, Louis Chan, Gregory Connor, Kent Daniel, Bernard Dumas, Ludger Hentschel, Philippe Henrotte, Ioulia Ioffe, Olivier Ledoit, Ludan Liu, Andrew Lo, Jesper Lund, Robert McDonald, Steve Ross, Jay Shanken, and seminar participants at Copenhagen Business School, HEC, INSEAD, MIT, Norwegian School of Management, Hong Kong University of Science and Technology, University of Rochester, University of Illinois at Urbana–Champaign, University of Utah, and the 2002 AFA meeting in Atlanta, and especially the referees, the editor, Rick Green, and Wayne Ferson for helpful comments. We are responsible for any errors or omissions.

the set of low beta stocks separately. The resulting two portfolios will have very low residual risk and different betas. Next, short the high beta portfolio and long the low beta portfolio to get rid of the systematic risk. Clearly, the second step involves taking extremely long and short positions when the dispersion in the assets' betas is small. That is why minimum variance portfolios take large positive and negative positions in the underlying assets.

Suppose efficient portfolios have extreme negative weights even in the population due to the presence of a single dominant factor. Then it would appear that imposing the nonnegativity constraints would lead to a loss in efficiency. However, empirical findings suggest that imposing these constraints improves the efficiency of optimal portfolios constructed using sample moments.[1]

In this paper, we show that imposing nonnegativity constraints on portfolio weights can help in practice even when Green and Hollifield are right—that is, the unobserved true covariance matrix is such that the minimum variance portfolio involves taking large negative positions in a number of assets even in the population. Consider the minimum variance portfolio constructed using the estimated covariance matrix with no-short-sales constraints in place. We show that each of the no-short-sales constraints is equivalent to reducing the estimated covariances of the corresponding asset with other assets by a certain amount. Stocks that have high covariances with other stocks tend to receive negative portfolio weights. Hence, to the extent that high estimated covariances are more likely to be caused by upward-biased estimation error, imposing nonnegativity constraints can reduce the sampling error. It follows from the theory of shrinkage estimators that imposing the no-short-sales constraint can help even when the constraints do not hold in the population.

We also study the impact of upper bounds on portfolio weights since they are also commonly imposed by practitioners. We show that each upper bound constraint is equivalent to increasing the sample covariances of the corresponding asset with other assets by a certain amount. Since stocks that have low covariances with other stocks tend to get extreme high portfolio weights, and these extreme low estimated covariances are more likely to be caused by downward-biased estimation error, this adjustment in estimated covariances could reduce sampling error and help the out-of-sample performance of the optimal portfolios. Our empirical evidence suggests that imposing upper bounds on portfolio weights does not lead to a significant improvement in the out-of-sample performance of minimum risk portfolios when no-short-sales restrictions are already in place.

It has long been recognized that mean-variance efficient portfolios constructed using sample means and sample covariance matrices perform poorly out of sample.[2] The primary reason is that the sample mean is an imprecise estimator of the population mean. The estimation error in the sample mean is so large that nothing much is lost in ignoring the mean altogether when no further information

[1] See Frost and Savarino (1988) for an excellent discussion.

[2] See Jobson and Korkie (1980, 1981), Frost and Savarino (1986, 1988), Jorion (1986), Michaud (1989), Best and Grauer (1991), and Black and Litterman (1992).

about the population mean is available. For example, the global minimum variance portfolio has as large an out-of-sample Sharpe ratio as other efficient portfolios when past historical average returns are used as proxies for expected returns.[3] In view of this we focus our attention on global minimum variance portfolios in this study.

We also examine the effect of portfolio weight constraints on the performance of the minimum tracking error portfolios. We are interested in minimum tracking error portfolios because it may be necessary in practice to construct portfolios using a subset of all available stocks that have low transactions costs and high liquidity to track certain benchmark indices that contain assets that are not actively traded.

In our study, we examine minimum variance portfolios constructed using 500 stocks. The covariance matrix of the returns has about 125,000 distinct parameters, whereas monthly returns for all traded stocks are only available for the past 900 months or so. This gives less than four degrees of freedom per estimated parameter. Consequently, the elements of the covariance matrix are estimated very imprecisely. Several solutions to this problem have been suggested. The first is to impose more structure on the covariance matrix to reduce the number of parameters that have to be estimated. This includes factor models and constant correlation models. The second approach is to use shrinkage estimators, shrinking the sample covariance matrix toward some target, such as the single-index model (Ledoit (1996, 1999)). The third approach is to use data of higher frequency, for example, to use daily data in place of monthly data. These three approaches are widely used both by practitioners and academics.

For any covariance matrix estimator, we can impose the usual portfolio weight constraints. According to Green and Hollifield (1992), these constraints are likely to be wrong in population and hence they introduce specification error. According to our analysis, these constraints can reduce sampling error. Therefore the gain from imposing these constraints depends on the trade-off between the reduction in sampling error and the increase in specification error. For covariance matrix estimators that have large sampling error, such as the sample covariance matrix, imposing these constraints is likely to be helpful. However, for the factor

[3] Using monthly stock index return data for G7 countries, Jorion (1985) convincingly argues that benefits from diversification are more likely to accrue from a reduction in risk. In the data set he examined, the global minimum variance portfolio had the best out-of-sample performance. It outperformed a classical tangent portfolio, the tangent portfolio constructed using the Bayes–Stein estimator for the vector of mean returns, and the value-weighted and equally weighted portfolios. Using simulation methods, Jorion (1986) shows that this conclusion is robust if the sample size is not large. Jorion (1991) finds that the minimum variance portfolio constructed using returns on seven industry stock index portfolios performed as well as the CRSP equally weighted and value-weighted stock indices during the January 1926 to December 1987 period, in out-of-sample tests. The performance was comparable to that of the tangent portfolio constructed using the Bayes–Stein estimator for the mean. Bloomfield, Leftwich, and Long (1977) find that portfolios constructed using mean-variance optimization did not dominate an equally weighted portfolio. On the other hand, Chan, Karceski, and Lakonishok (1999) document that the constrained global minimum variance portfolios outperform the equally weighted portfolio.

models and shrinkage estimators, imposing such constraint is likely to hurt. This is what we find empirically. Sampling errors are likely to be less important when the assets we use to construct minimum variance portfolios are themselves large portfolios. In that case, imposing nonnegativity constraints should not help much. Again, this is what we find.

Our main empirical findings are the following:

1. For factor models and shrinkage estimators, constraining portfolio weights to be nonnegative results in a slight reduction in performance. In contrast when the nonnegativity constraints on portfolio weights are in place, minimum variance and minimum tracking error portfolios constructed using the sample covariance matrix perform almost as well as those constructed using factor models, shrinkage estimators, or daily returns.[4]
2. When short sales are allowed, minimum variance portfolios and minimum tracking error portfolios constructed using the daily return sample covariance matrix performs the best. When covariance matrices are estimated using daily data, corrections for microstructure effects that have been suggested in the literature do not lead to superior performance. So far, this has not been recognized in the literature.
3. Tangency portfolios, whether constrained or not, do not perform as well as the global minimum variance portfolios in terms of the out-of-sample Sharpe ratio. This means that the estimates of the mean returns are so noisy that simply imposing the portfolio weight constraint is not enough, even though the constraints still have a shrinkage effect.
4. Monte Carlo simulations indicate that when nonnegativity constraints are in place, global minimum variance portfolios constructed using the sample covariance matrix can perform just as well as the covariance matrix estimators constructed by imposing the factor structure, even when returns do have a dominant factor structure in the population.

The rest of the paper is organized as follows. In Section I, we provide a theoretical analysis of the shrinkage-like effect of imposing the no-short-sales restriction and upper bounds on portfolio weights when constructing global minimum risk portfolios. In Section II, we use simulation to evaluate the trade-off between specification error and sampling error. This trade-off clearly depends on the true covariance structure of the assets, which is not observed. We therefore calibrate the covariance structure in the simulation to that of the U.S. stocks. Using the simulation, we provide some guidance as to when nonnegativity constraints start to hurt. In Section III, we empirically examine the effects of portfolio weight constraints using the out-of-sample performances of the constrained and unconstrained optimal portfolios. We also briefly discuss the role of portfolio weight

[4] It is well recognized in the literature that imposing portfolio weight constraints leads to superior out-of-sample performance of mean-variance efficient portfolios. However, to our knowledge, no one has noticed that the performance improvement is so large that it is comparable to that attained using the other alternatives.

constraints in constructing tangency portfolios. We conclude the paper in Section IV. Proofs and details about the construction of various covariance matrix estimates using daily returns are given in the Appendices.

I. The Effect of Portfolio Weight Constraints

A. Some Theoretical Results

Given an estimated covariance matrix S, the global portfolio variance minimization problem when portfolio weights are constrained to satisfy both a lower bound of zero and an upper bound of $\bar{\omega}$ is given by:

$$\min_{\omega} \ \omega' S \omega \tag{1}$$

$$\text{s.t.} \ \sum_{\mathrm{i}} \omega_{\mathrm{i}} = 1 \tag{2}$$

$$\omega_i \geq 0, \ i = 1, 2, \ldots, N. \tag{3}$$

$$\omega_i \leq \bar{\omega}, \ i = 1, 2, \ldots, N. \tag{4}$$

The Kuhn–Tucker conditions (necessary and sufficient) are

$$\sum_j S_{i,j}\omega_j - \lambda_i + \delta_i \ = \ \lambda_0 \geq 0, \quad i = 1, 2, \ldots, N. \tag{5}$$

$$\lambda_i \geq 0, \ \text{ and } \lambda_i = 0 \text{ if } \omega_i > 0, \quad i = 1, 2, \ldots, N. \tag{6}$$

$$\delta_i \geq 0, \ \text{ and } \delta_i = 0 \text{ if } \omega_i < \bar{\omega}, \quad i = 1, 2, \ldots, N. \tag{7}$$

Here $\lambda = (\lambda_1, \ldots, \lambda_N)'$ are the Lagrange multipliers for the nonnegativity constraints (3), $\delta = (\delta_1, \ldots, \delta_N)'$ the multipliers for the constraints (4), and λ_0 is the multiplier for (2).

Denote a solution to the constrained portfolio variance minimization problem (1)–(4) as ω^{++} (S). Let $\mathbf{1}$ denote the column vector of ones. Then we have the following proposition.

PROPOSITION 1: *Let*

$$\tilde{S} = S + (\delta \mathbf{1}' + \mathbf{1}\delta') - (\lambda \mathbf{1}' + \mathbf{1}\lambda'). \tag{8}$$

Then $\tilde{S}$ is symmetric and positive semi-definite, and ω^{++} (S) is one of its global minimum variance portfolios.

All proofs are given in Appendix A.

This result shows that constructing a constrained global minimum variance portfolio from S is equivalent to constructing a (unconstrained) minimum variance portfolio from $\tilde{S} = S + (\delta \mathbf{1}' + \mathbf{1}\delta') - (\lambda \mathbf{1}' + \mathbf{1}\lambda')$. Later we will interpret $\tilde{S}$ as a shrunk version of S, and argue that this shrinkage can reduce sampling error.

In general, given a constrained optimal portfolio $\omega^{++}(S)$, there are many covariance matrix estimates that have $\omega^{++}(S)$ as their (unconstrained) minimum variance portfolio. Is there anything special about $\tilde{S}$? We do have an answer to this question when returns are jointly normal and S is the MLE of the population covariance matrix.

Let the $N \times 1$ return vector $h_t = (r_{1t}, r_{2t}, \ldots, r_{Nt})'$ be i.i.d. normal $N(\mu, \Omega)$. Then the MLE of Ω is $S = \sum_{t=1}^{T} (h_t - \bar{h})(h_t - \bar{h})'/T$. The likelihood function depends on both μ and Ω, even though we want to estimate Ω only. To get rid of the dependence on μ, recall that for any estimate of the covariance matrix, the MLE of the mean is always the sample mean (Morrison (1990)). With this estimate of the mean, the log-likelihood (as a function of the covariance matrix alone) becomes

$$l(\Omega) = \mathrm{CONST} - \frac{T}{2} \ln |\Omega| - \frac{T}{2} \operatorname{tr}\left(S\Omega^{-1}\right). \tag{9}$$

This can also be considered as the likelihood function of Ω^{-1} and is defined for nonsingular Ω.

Now consider the constrained MLE of Ω, subject to the constraint that the global minimum variance portfolio constructed from Ω satisfies the weight constraints (3)–(4). Let $\Omega_{i,j}$ denote the (i, j)-th element of Ω and $\Omega^{i,j}$ denote the (i, j)-th element of Ω^{-1}; then the constraints are

$$\sum_j \Omega^{i,j} \ \geq \ 0, \ i = 1, 2, \ldots, N. \tag{10}$$

$$\sum_j \Omega^{i,j} \ \leq \ \bar{\omega} \sum_k \sum_j \Omega^{k,j}. \tag{11}$$

So the constrained maximum likelihood (ML) problem is to maximize (9), subject to constraints (10)–(11). We have the following proposition.

Proposition 2: *Assume that returns are jointly i.i.d. normal $N(\mu, \Omega)$. Let S be the unconstrained MLE of Ω.*

1. *Given S, let $\{\lambda_i, \delta_i, \omega_i\}_{i=1,\ldots,N}$ be a solution to the constrained portfolio variance minimization problem (1)–(4), and construct $\tilde{S}$ according to (8). Assume $\tilde{S}$ is nonsingular. Then, $\tilde{S}$ and $\{\lambda_i(1 - \bar{\omega})\delta_i\}_{i=1,\ldots,N}$ jointly satisfy the first-order conditions for the constrained ML problem.*
2. *Let $\tilde{S}$ and $\{\lambda_i, \delta_i\}_{i=1,\ldots,N}$ jointly satisfy the first-order conditions for the constrained ML problem. For $i = 1, \ldots, N$, define $= \omega_i \sum_j \tilde{S}^{i,j} / \sum_k \tilde{S}^{k,l}$, the normalized row sums of $\tilde{S}^{-1}$. Then $\{\lambda_i, \delta_i/(1 - \bar{\omega}), \omega_i\}_{i=1,\ldots,N}$ is a solution to the constrained portfolio variance minimization problem (1)–(4), given S.*

According to this proposition the $\tilde{S}$ constructed from the solution to the constrained global variance minimization problem is the ML estimator of the covariance matrix, subject to the condition that the global minimum variance portfolio weights satisfy the nonnegativity and upper bound constraints. So we could impose the constraints in the estimation stage instead of the optimization stage and the result would be the same.

When only the nonnegativity constraint is imposed, the vector of Lagrange multipliers for the upper bound will be zero. So $\tilde{S} = S - (\lambda \mathbf{1}' + \mathbf{1}\lambda')$, and we can simplify the statements in Proposition 2 in a straightforward way.

B. A Shrinkage Interpretation of the Effects of Portfolio Weight Constraints

Let us first examine the effect of the nonnegativity constraint. Consider the unconstrained global portfolio variance minimization. The first order condition is

$$\sum_{j=1}^{N} \omega_j S_{i,j} = \lambda_0 \geq 0,\ i = 1, 2, \ldots, N. \tag{12}$$

The above condition says that at the optimum, stock *i*'s marginal contribution to the portfolio variance is the same as stock *j*'s, for any *i* and *j*.

Suppose stock *i* tends to have higher covariances with other stocks, that is, the *i*th row of S tends to have larger elements than other rows. Then stock *i*'s marginal contribution to the portfolio variance will tend to be larger than the marginal contributions of other stocks. Therefore, to achieve optimality, we need to reduce stock *i*'s portfolio weight. Stock *i* may even have negative weight if its covariances with other stocks are sufficiently high. Therefore, a stock tends to receive a negative portfolio weight in the global minimum variance portfolio if it has higher variance and higher covariances with other stocks. In a one-factor structure, these are the high-beta stocks.

With only the nonnegativity constraint, Proposition 1 implies that constructing the constrained global minimum variance portfolio from S is equivalent to constructing the (unconstrained) global minimum variance portfolio from $\tilde{S} = S - (\mathbf{1}'\lambda + \lambda\mathbf{1}')$. Notice that the effect of imposing the constraint is that whenever the nonnegativity constraint is binding for stock i, its covariances with other stocks, S_{ij}, $j \neq i$, are reduced by $\lambda_i + \lambda_j$ a positive quantity, and its variance is reduced by $2\lambda_i$. Therefore, the new covariance matrix estimate, $\tilde{S}$, is constructed by shrinking the large covariances that would otherwise imply negative weights. Since the largest covariance estimates are more likely caused by upward-biased estimation error, this shrinking may reduce the estimation error.

We can interpret the effect of the upper bounds on portfolio weights in a similar way. In the unconstrained portfolio variance minimization problem, those stocks with low covariances with other stocks tend to receive high portfolio weights. When the upper bound is imposed, it tends to be binding for these stocks. According to Proposition 1, constructing the upper-bound constrained minimum variance portfolio using S is the same as constructing the unconstrained minimum variance portfolio from $\tilde{S} = S + (\mathbf{1}\delta' + \delta\mathbf{1}')$. The effect of imposing the constraint is that, whenever the upper bound is binding for stock i, its variance is raised by $2\delta_i$ and its covariance with another stock j is increased by $(\delta_i + \delta_j)$. Since the upper bound tends to be binding for those stocks with low covariances with other stocks, and low estimated covariances are more likely to be plagued by downward-biased estimation error, this adjustment may reduce sampling error and achieve a shrinkage-like effect.

We now illustrate this shrinkage-like effect of portfolio weight constraints through an example. The covariance matrix estimate S is the sample covariance matrix of a random sample of 500 stocks among all domestic common stocks that have all monthly returns from 1990 to 2000 (132 months). We consider the effect of imposing the nonnegativity constraint in the global portfolio variance minimization problem. (The effect of upper bounds on portfolio weights is qualitatively similar and hence is not reported.) Panel A of Table I reports summary statistics for the row average covariances of S, the Lagrange multipliers for the nonnegativity constraint, and the row average covariances of the adjusted estimate $\tilde{S}$. The row average covariances of S are the row sums of S divided by 500. Recall that the effect of the nonnegativity constraint is to reduce each element in the ith row of S by λ_i. Therefore we report row average covariances instead of row sums of S, since the former can be compared with the Lagrange multipliers to give a sense of the magnitude of the adjustment due to the constraint.

We can see that the row average covariances of S vary over a wide range: The lowest row average covariance is -0.27×10^{-4}, and the highest is 51.86×10^{-4}. Thus the range (i.e., the highest minus the lowest row average covariance) is 52.13×10^{-4}.

The row average covariances of the adjusted covariance estimate $\tilde{S}$ are less variable. Their range is only 44.42×10^{-4}. So those rows that have the largest covariances tend to be reduced due to the nonnegativity constraint. This is consistent with our previous analysis.

For comparison purposes, we also reported the row average covariances of the Ledoit estimate of the covariance matrix S_{Ledoit}, which is a shrinkage estimator of the covariance matrix (Ledoit (1996, 1999)). We can see that the mean, standard deviation, minimum, and maximum of the row averages of $\tilde{S}$ are very similar to those of S_{Ledoit}. This is consistent with the interpretation that imposing the nonnegativity constraint has a shrinkage-like effect.

If imposing the nonnegativity constraint is similar to shrinking S toward the one-factor model as in Ledoit's estimator, then we expect changes in row averages from S to $\tilde{S}$ to be correlated with changes in row averages from S to S_{Ledoit}. Panel B of Table I examines this issue.

We first regress the changes in row averages from S to $\tilde{S}$ on the row averages of S, yielding a highly significant coefficient of -0.143. We then regress the changes in row averages from S to S_{Ledoit} on the row averages of S, obtaining a highly significant coefficient of -0.226. Finally we regress the changes in row averages from S to $\tilde{S}$ on the changes in row averages from S to S_{Ledoit}, yielding a highly significant coefficient of 0.296.

The first regression says that, from S to $\tilde{S}$ the highest row averages tend to be reduced the most. This is consistent with our previous interpretation of the effect of the nonnegativity constraint. The last regression shows that the effect of imposing the nonnegativity constraint on the row averages of S is to some extent similar to the effect of shrinking S toward the one-factor model.

Table I

The Effect of Portfolio Weight Constraints

This table shows the effects of nonnegativity constraints on portfolio weights. The covariance matrix S is the sample covariance matrix of monthly returns of a random sample of 500 stocks from 1990 to 2000 (132 months). The variable λ is the vector of Lagrange multipliers on the nonnegativity constraint. The matrix $\tilde{S}$ is our shrunk version of S; S_{Ledoit} is the Ledoit estimate of the covariance matrix. The numbers below the regression coefficients are t-statistics.

Panel A: Summary statistics of the row average covariances of S, unconstrained portfolio weights, and the Lagrange multipliers (all numbers are scaled by a factor of 10^4.)

Variable	Mean	Std. Dev	Minimum	Maximum	Max − Min
Row average covariances of S	17.95	9.39	− 0.27	51.86	52.13
$\lambda_i s$	2.31	1.87	0	7.93	7.93
Row average covariances of $\tilde{S}$	13.32	8.15	− 2.59	41.82	44.41
Row average covariances of S_{Ledoit}	14.34	7.60	− 3.15	42.21	45.36

Panel B: Regressions of changes in row averages from S to $\tilde{S}$ and from S to S_{Ledoit} (the intercept terms are omitted for brevity)

Independent Variable	Row Averages of S	Row Averages (S_{Ledoit}–S)	Adjusted R^2	F-value (Significance)
Row averages of ($\tilde{S}$–S)	− 0.143		0.514	527.7
	(− 22.97)			(0.000)
Row averages of (S_{Ledoit} − S)	− 0.226		0.469	442.5
	(− 21.04)			(0.000)
Row averages of ($\tilde{S}$–S)		0.296	0.239	157.6
		(12.55)		(0.000)

Panel C: Regressions of the changes in the distinct elements from S to $\tilde{S}$ and from S to S_{Ledoit} (the intercept terms are omitted for brevity)

Independent Variable	$S_{ij, j \geq i}$	$(S_{Ledoit}–S)_{ij, j \geq i}$	Adj. R^2	F-value (Significance)
$(\tilde{S} - S)_{ij, j \geq i}$	− 0.0271		0.099	13810
	(− 117.5)			(0.000)
$(S_{Ledoit}–S)_{ij, j \geq i}$	− 0.336		0.536	144599
	(− 380.3)			(0.000)
$(\tilde{S} - S)_{ij, j \geq i}$		0.0283(54.2)	0.023	2933.4
		(54.2)		(0.000)

Panel C of Table I is similar to Panel B, except that we examine the changes in distinct elements from S to $\tilde{S}$. The results are qualitatively similar to those in Panel B. Note that even though some of the adjusted R^2 appear low in Panel C, the regressions are highly significant as indicated by the F-values and t-statistics.

To summarize, Table I shows that the effect of imposing the nonnegativity constraint is similar to shrinking S toward the one-factor model. This is consistent with a shrinkage interpretation of the nonnegativity constraint.

We would like to emphasize that we are not saying that $\tilde{S}$ is an equally good or better estimator than S_{Ledoit}. Our interest is not to derive a better covariance matrix estimator. Instead, we are trying to explain why imposing the constraints that are wrong in population helps the out-of-sample performance of the optimal portfolios in practice. We do that by showing that the constraints have a shrinkage-like effect on the covariance matrix estimate.

C. The Role of Constraints in Constructing Mean-Variance Efficient Portfolios

Under certain conditions, we can also give a shrinkage-like interpretation to the constraints in constructing mean-variance efficient portfolios. Liu (2001) and an earlier version of this paper (Jagannathan and Ma (2002)) give more details. When S is nonsingular, the effect of the lower bound on portfolio weights is to adjust the mean returns upward by an amount proportional to the corresponding Lagrange multiplier, and the effect of the upper bound on portfolio weights is a similar downward adjustment.

From an empirical point of view, this result is less interesting because tangency portfolios even with portfolio weight constraints perform worse than the global minimum variance portfolios out of sample. It is therefore necessary to bring additional information about the population mean instead of relying on sample mean as an estimator of population mean as in Black and Litterman (1992), Pastor and Stambaugh (2000), and Wang (2002).

II. A Simulation Study of the Trade-off Between Specification Error and Sampling Error

In the last section, we argued that imposing the nonnegativity constraint in global portfolio variance minimization has a shrinkage-like effect. This shrinkage-like effect can reduce sampling error. On the other hand, as Green and Hollifield (1992) point out, the constraint will not, in general, hold in the population when asset returns have a single dominant factor. Whether the nonnegativity constraint would help depends on the trade-off between these two effects: the reduction in sampling error on the one hand and the increase in specification error on the other hand. That trade-off would depend on, among other things, the true covariance structure. Since we do not know the true covariance structure, we cannot give a quantitative answer that holds true for all underlying covariance structures in the population. We therefore rely on using Monte Carlo simulation in this section to further examine the trade-off between the two effects. By comparing the simulation results with the results using the real data, which are reported in the next section, we are also able to provide additional insights regarding the true covariance structure of U.S. stock returns.

We assume a two-factor return structure in the simulation.[5] We consider a two-factor structure because the literature often finds more than one factor in the

[5] We also performed a simulation with a one-factor structure. The results are qualitatively similar. These results are available upon request.

return structure of U.S. stocks. Without loss of generality, we set the factor variances to unity and the covariance between them to zero. We draw the betas for the first factor for the stocks from an i.i.d. Normal distribution with mean one and standard deviation σ_β, which varies from 0 to 0.4 across different sets of simulations. The stocks' betas with respect to the second factor are drawn from an i.i.d. Normal distribution with mean zero and standard deviation 0.2 when σ_β, is nonzero. Otherwise the second factor is the cross-sectional standard deviation of betas with respect to/also set to zero. The residual variances are set to be constant over time, but cross-sectionally, they are drawn from an i.i.d. lognormal (0.8, 0.7) distribution. When $\sigma_\beta = 0.4$, the distribution of the betas with respect to the two factors roughly matches the empirical distribution of the first two betas for the NYSE stocks, with the two factors being the first two of the five Connor–Korajczyk factors. Furthermore, in this case, the ratio of residual variance to total variance matches the same ratio in the NYSE stocks.[6]

We allow the number of stocks, N, to range from 30 to 300. For each specification of N and σ_β, we draw the betas and the residual variances for the stocks as specified above. These betas and residual variances are fixed for this specification even as we change the sample size T.

Given the sampled betas and residual variances, we then draw returns of different sample sizes. We first set $T = 60$ and then vary it from $N + 30$ to $N + 210$ by an increment of 60. We allow $T = 60$ because this is the sample size used in the simulation with real data in the next section, and having the same sample size allows us to compare the results.

Given the simulated returns data, we estimate the return covariance matrix, and form the nonnegativity constrained and unconstrained global minimum variance portfolios. We then calculate the ex post variances of these portfolios using the assumed true covariance matrix. This procedure is repeated 10 times, and the averages of ex post variances are reported.

We consider two covariance matrix estimators. The first is the sample covariance matrix. The second is the one-factor model. When the covariance matrix is estimated with the one-factor model, we assume the factor return is known, and is the return of the first factor used in the simulation.

Before we discuss the effect of imposing the nonnegativity constraint, we refer the reader to Table II, which gives the short interest in the unconstrained global minimum variance portfolios constructed from the true covariance matrix. This table gives us a measure about the extent to which the nonnegativity constraint is violated (i.e., the specification error of the nonnegativity constraint for each covariance matrix structure). Two patterns emerge: (1) For a given dispersion in the betas, as the number of assets increases, the nonnegativity constraint becomes more severely violated in population. (2) When the number of assets is not too

[6] When we scale the factor variances and residual variances up or down by a common factor, the population and sample covariance matrices are also scaled up or down by the same factor, and hence optimal portfolio weights will stay the same, as will the relative performance of different portfolios. Therefore, the discussion below is not affected by the fact that we normalized the factor variances to one.

Table II

Short Interests in the Global Minimum Variance Portfolios

The true return covariance has a two-factor structure. Across stocks, the betas with respect to the first factor are Normally distributed with mean one and standard deviation δ_β, the betas with respect to the second factor are Normally distributed with mean zero and standard deviation 0.2. The two factor returns have unit variances and zero covariance. The stocks' residual variances are constant over time, but follow a Lognormal(0.8, 0.7) distribution cross-sectionally. The number of stocks is N. For each (δ_β, N) pair, we draw the betas and the residual variances according to the above-mentioned distributional assumptions, and then calculate the true stock return covariance matrix. We then calculate the unconstrained global minimum variance portfolio and its short interest (i.e., the total short positions). This procedure is repeated 10 times and the averages are reported below. The short interests give the extent to which the nonnegativity constraint is wrong.

	σ_β			
N	0.1	0.2	0.3	0.4
30	− 0.181	− 0.401	− 0.391	− 0.350
60	− 0.554	− 0.706	− 0.584	− 0.468
120	− 1.188	− 1.012	− 0.737	− 0.515
180	− 1.578	− 1.161	− 0.828	− 0.555
240	− 1.845	− 1.250	− 0.816	− 0.534
300	− 2.042	− 1.308	− 0.831	− 0.526

small, the smaller the dispersion in the betas (while there is some dispersion), the more severely the nonnegativity constraint is violated. Both patterns can be understood by following Green and Hollifield's (1992) two-step procedure in forming the global minimum variance portfolios.

Table III presents a subset of our simulation results.[7] It compares the effects of imposing the nonnegativity constraint with that of the single-factor constraint. The table reports the percentage reductions in the ex post standard deviations of the global minimum variance portfolios constructed by imposing these constraints relative to that of the unconstrained global minimum variance portfolios constructed from the sample covariance matrix, and relative to the equal-weighted portfolio. The three panels, $\sigma_\beta = 0$, $\sigma_\beta = 0.2$, and $\sigma_\beta = 0.4$, represent the cases where the nonnegativity constraint is correct, is severely wrong, and is moderately wrong, respectively. (When there is no dispersion in the betas, the nonnegativity constraint holds in population.) We report the cases where the number of stocks is either 30 or 300, and the sample size is either 60 or 360.

Panel A shows that when the nonnegativity constraint is correct, imposing the nonnegativity constraint always helps. The gain from imposing the nonnegativity constraint is comparable to that obtained by imposing the one-factor structure. In fact, this is true for all the cases in our simulation, not only those cases reported in the table.

[7] To save space, we do not report the full result of the simulations, which is available upon request.

Table III

Effects of Imposing Nonnegativity Constraints and the Single-Factor Constraint

This table shows the percentage reduction in the ex post standard deviations of the global minimum variance portfolios when the nonnegativity and/or single-factor model constraints are imposed, relative to the ex post variances of the global minimum variance portfolios constructed from the sample covariance matrix and the equal-weighted portfolio. N is the number of assets; T is the sample size used to estimate the covariance matrix. The value σ_β is the cross-sectional standard deviation in the stocks' betas with respect to the first factor. When σ_β is 0, the nonnegativity constraint is correct in population; when σ_β is 0.4, the nonnegativity constraint is violated roughly to the same extent as using the NYSE stocks; and when σ_β is 0.2, the nonnegativity constraint is more severely violated than when σ_β is 0.4. The rows with T = Infinity are the population results.

		Percentage Reduction in ex post Standard Deviation of the Optimal Portfolios from Imposing the Constraints					
		Relative to the Portfolio Constructed from the Sample Covariance Matrix			Relative to the Equal-Weighted Portfolio		
N	T	Single Factor Model	Nonnegativity Constraint	Both	Single Factor	Constraint Constraint	Both
			Panel A: $\sigma_\beta = 0$				
30	60	21.6	25.2	26.0	− 12.2	− 7.0	− 6.0
30	360	0.8	1.2	1.4	− 1.7	− 1.3	− 1.0
30	Infinity	0.0	0.0	0.0	1.1	1.1	1.1
300	60	—	—	—	− 8.1	− 11.6	− 6.9
300	360	54.6	58.8	59.1	− 14.6	− 3.9	− 3.2
300	Infinity	0.0	0.0	0.0	0.2	0.2	0.2
			Panel B: $\sigma_\beta = 0.2$				
30	60	20.7	16.2	18.7	8.4	3.3	6.1
30	360	1.7	− 3.2	− 2.6	14.8	10.6	11.1
30	Infinity	− 0.3	− 5.6	− 5.8	16.8	12.4	12.3
300	60	—	—	—	45.9	18.1	23.2
300	360	56.2	26.1	26.5	60.3	33.0	33.3
300	Infinity	− 5.0	− 80.7	− 81.3	62.8	35.9	35.7
			Panel C: $\sigma_\beta = 0.4$				
30	60	27.2	16.8	21.2	38.9	30.2	33.9
30	360	2.3	− 7.6	− 7.1	42.6	36.8	37.0
30	Infinity	− 0.1	− 10.0	− 10.0	43.3	37.7	37.7
300	60	—	—	—	72.9	50.8	57.6
300	360	59.2	25.0	27.2	79.3	61.9	63.0
300	Infinity	− 0.8	− 79.4	− 80.7	79.9	64.2	64.0

Across all three panels, the numbers in the last two columns are comparable. This means that if, for some reason, the nonnegativity constraint has to be imposed, then whether we use the sample covariance matrix or the one-factor model to estimate the covariance matrix makes little difference. Again this is true not only for the cases in Table III but for all the cases in our simulation.

When $\sigma_\beta > 0$, so that the true global minimum variance portfolio has negative weights, the nonnegativity constraint always hurts out-of-sample performance if we use the one-factor model to estimate the covariance matrix, and the deterioration in performance is dramatic if the number of stocks is large. However, imposing the nonnegativity constraint helps when the sample covariance matrix is used and there are too many stocks relative to the length of the time series of observations used to estimate the covariance matrix. Again this is true in general and not only for the cases reported in the table.

The numbers in the last two columns indicate that, paradoxically, the nonnegativity-constrained portfolios do worse than the equal-weighted index when the nonnegativity constraint is correct (i.e., Panel A), and do better than the equal-weighted index when the nonnegativity constraint is wrong (i.e., Panels B and C). This is because when there is dispersion in the betas (so that the nonnegativity constraint is wrong in population), even if short sales are not allowed, we can still reduce the exposure to systematic risk by investing only in the low-beta stocks. Thus we can achieve lower risk than in the equal-weighted index. On the other hand, when there is no dispersion in the betas (so that the nonnegativity constraint is correct in population), we cannot reduce systematic risk at all. Hence there is not much difference between optimal diversification and naive diversification, even though we know the population covariance matrix. So optimal diversification based on sample estimates can perform worse than naive diversification due to sampling error.

Based on the results from the full set of simulation, we also find a meaningful trade-off between specification error and sampling error when $\sigma_\beta > 0$ and the covariance matrix estimate is the sample covariance matrix. First, when $\sigma_\beta = 0.1$, the nonnegativity constraint always improves the out-of-sample performance. On the other hand, as σ_β goes from 0.2 to 0.4, we always find cutoff points for the sample size T, so that when we have fewer observations than the cutoff points, the nonnegativity constraint improves the out-of-sample performance, and when we have more observations than the cutoff points, the nonnegativity constraint hurts the out-of-sample performance. These cutoff points grow as we increase the number of assets (as expected), but are not sensitive to changes in σ_β. These cutoff points are presented in Table IV. Recall that $\sigma_\beta = 0.4$ roughly corresponds to the covariance structure of the NYSE stocks, and that $\sigma_\beta = 0.2$ or 0.3 represents cases where the nonnegativity constraint is more severely violated than when $\sigma_\beta = 0.4$. Since these cutoff points are not sensitive to changes in σ_β in this range, we can treat them as a rough guideline for practitioners who use U.S. stocks.

III. The Effect of Portfolio Weight Constraints: An Empirical Examination

A. Data and Methodology

In this section, we examine empirically the effect of portfolio weight constraints. As we have said before, whether these weight constraints help or

Table IV

Cutoff Points for the Sample Sizes When the Nonnegativity Constraints Start to Hurt the Out-of-Sample Performance

The true return covariance has a two-factor structure. Across stocks, the betas with respect to the first factor are normally distributed with mean one and standard deviation σ_β. The number of stocks is N, and the sample size is T. We estimate the covariance matrix for the stocks using simulated returns of sample size T using the sample covariance matrix. Then we form the nonnegativity constrained and unconstrained global minimum variance portfolios using the estimated covariance matrix, and calculate the ex post variances of these portfolios. For each (δ_β, N) pair, we report the cutoff point for T, such that when the sample size is greater than or equal to the cutoff point, the nonnegativity constraint starts to hurt the out-of-sample performance. An entry of na means such cutoff point is not found in that set of simulation.

	σ_β		
N	0.2	0.3	0.4
30	na	120	120
60	150	150	150
120	270	270	330
180	330	270	330
240	390	390	450
300	450	450	450

hurt is an empirical issue and depends on the specific covariance matrix estimator. For the estimators that have large sampling errors, such as the sample covariance matrix, the portfolio weight constraints are likely to be helpful, as documented by Frost and Savarino (1988) and demonstrated again in the simulation study in the last section. However, for the factor models and shrinkage estimators, the portfolio weight constraints are likely to be harmful.

We examine the effect of the portfolio weight constraints on the out-of-sample performance of minimum variance and minimum tracking error portfolios formed using a number of covariance matrix estimators. This is done following the methodology in Chan et al. (1999). At the end of each April from 1968 to 1998, we randomly choose 500 stocks from all common domestic stocks traded on the NYSE and the AMEX, with stock price greater than five dollars, market capitalization more than the 80th percentile of the size distribution of NYSE firms, and with monthly return data for all the immediately preceding 60 months. We use return data for the preceding 60 months to estimate the covariance matrix of the returns of the 500 stocks. For estimators that use daily data, the daily returns during the previous 60 months of the same 500 stocks are used. When a daily return is missing, the equally weighted market return of that day is used instead.

When variance minimization is the objective, we form three global minimum variance portfolios using each covariance matrix estimator—only two if the covariance matrix estimate is singular. The first portfolio is constructed without imposing any restrictions on portfolio weights, the second is subject to the constraint that portfolio weights should be nonnegative, 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 is held for one year. Their monthly returns are recorded, and at the end of April of the next year, the same process is repeated. This gives at most three minimum variance portfolios that have postformation monthly returns from May 1968 to April 1999 for each covariance matrix estimator. We use the standard deviation of the monthly returns on these portfolios to compare the different covariance matrix estimators.

For tracking error minimization, following Chan et al. (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 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.

B. Covariance Matrix Estimators

The first estimator is the sample covariance matrix:

$$S_N = \frac{1}{T-1}\sum_{t=1}^{T}\left(h_t - \bar{h}\right)\left(h_t - \bar{h}\right),$$

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

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

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

$$S_1 = s_m^2 BB' + D. \tag{13}$$

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 optimal shrinkage estimator of Ledoit (1999). It is a weighted average of the sample covariance matrix and the one-factor model-based estimator:

$$S_L = \frac{\alpha}{T} S_1 + \left(1 - \frac{\alpha}{T}\right) S_N,$$

where α is a parameter that determines the shrinkage intensity that is estimated from the data. Ledoit (1999) shows this estimator outperforms the constant correlation model (Elton and Gruber (1973) and Schwert and Seguin (1990)), the single-factor model, the industry factor model, and the principal component model with five principal components.[8] For the fourth set of estimators, we consider the Fama and French (1993) three-factor model, the Connor and Korajczyk (1986, 1988) five-factor model, and a three-factor version of it which includes only the first three of the five factors. This gives three additional covariance matrix estimators, each corresponding to one of these multifactor models.[9]

Finally we consider several covariance matrix estimators that use daily return data. These include the daily return sample covariance matrix, daily one-factor model, daily Fama–French three-factor model, and daily Connor–Korajczyk five-factor and three-factor models. These models are similar to the corresponding monthly models. However, we incorporate the corrections for microstructure effects suggested by Scholes and Williams (1977), Dimson (1979), and Cohen et al. (1983). We also develop a new estimator for the covariance matrix of returns using daily return data that nests these three estimators. Details on the estimation of the monthly return covariance matrix using daily return data that allow for microstructure effects are provided in Appendix B.

C. Empirical Results

Table V gives the characteristics of minimum variance portfolios constructed using various covariance matrix estimates.[10] Judging by the ex post standard deviations of the optimal portfolios, the shrinkage estimator proposed by Ledoit (1999) and the sample covariance matrix of daily return are the best performers. Since the Ledoit estimator is a particular weighted average of the one-factor model and the sample covariance matrix, we examined whether a simple average of the two estimators would do equally well. A randomly weighted average of the one-factor model and the sample covariance matrix has an annualized out-of-sample standard deviation of 10.34 percentage points (not reported in Table V), which is not much different from the 10.76 percentage points for the Ledoit estimator. We thus suspect that the sampling errors associated with the estimated optimal shrinkage intensity in the Ledoit estimator is rather large. Notice that even for these two best estimators, the unconstrained global minimum variance portfolios involve taking short positions that are over 80 percent of the portfolio value.

Next, we turn to the case where no-short-sale restrictions are imposed. Surprisingly, the minimum variance portfolio constructed using the monthly sample covariance matrix compares favorably with all the other covariance matrix esti-

[8] For tracking error variance minimization, we also considered Ledoit's (1996) estimator that shrinks the sample covariance matrix toward the identity matrix. The results are similar and we do not report them.

[9] For tracking error variance minimization, the loadings on the first factor in these multifactor models are set to zero for every stock.

[10] To save space, we only report a subset of the simulation results in Tables V to VII. The results of using Connor–Korajczyk three- and five-factor models are very similar to those using the Fama–French three-factor model. This is true when daily returns are used as well. So these results are not reported. Several versions of the daily one-factor model are dominated by the daily return sample covariance matrix. So these results are not reported either.

Table V

Ex Post Mean, Standard Deviation, and Other Characteristics of the Global Minimum Variance Portfolios

At the end of April each year from 1968 to 1997, the covariance matrix of a random sample of 500 stocks is estimated according to various estimators. We use these covariance matrix estimates to construct the global minimum variance portfolios, both constrained and unconstrained. We hold the portfolios for the next 12 months and their monthly returns are recorded. The ex post means, standard deviations, and other characteristics of these portfolios are reported. The C after an estimator indicates the nonnegativity constrained portfolios, and D after an estimator indicates the portfolio with both the nonnegativity constraint and an upper bound of two percent. For the equal-weighted and value-weighted portfolios of 25 stocks, the 25 stocks are randomly selected from the 500. Means and standard deviations are in percentage per year; maximum weight, minimum weight, and short interest numbers are in percentage.

Covariance matrix estimator	Mean	Std Dev	Max. weight	Min. weight	Short Interest	No. of Positive Weights
Panel A: Covariance matrix estimated using monthly return data						
Sample covariance matrix, C	13.55	12.43	18.4	0	0	24.1
Sample covariance matrix, D	13.55	12.85	2.0	0	0	59.5
One-factor model	13.99	11.69	3.8	− 0.9	− 50.7	268.8
One-factor model, C	12.68	12.62	11.3	0	0	39.3
One-factor model, D	13.51	12.50	2.0	0	0	63.1
Ledoit	13.09	10.76	4.9	− 1.7	− 80.6	283.2
Ledoit, C	12.79	12.29	13.4	0	0	39.7
Ledoit, D	13.49	12.43	2.0	0	0	65.4
Fama–French 3-factor model	13.04	11.35	4.2	− 1.4	− 63.5	284.1
Fama–French 3-factor model, C	12.65	12.38	12.2	0	0	40.1
Fama–French 3-factor model, D	13.34	12.53	2.0	0	0	64.0
Panel B: Covariance matrix estimated using daily return data						
Sample covariance matrix	14.06	10.64	6.3	− 2.5	− 122.4	270.7
Sample covariance matrix, C	13.95	12.34	8.8	0	0	64.7
Sample covariance matrix, D	14.22	12.28	2.0	0	0	81.1
Sample covariance matrix of CHMSW, C	13.89	12.31	9.1	0	0	62.5
Sample covariance matrix of CHMSW, D	14.18	12.25	2.0	0	0	80.1
Sample covariance matrix (New)	13.94	10.60	6.3	− 2.6	− 128.5	269.5
Sample covariance matrix (New), C	13.81	12.26	9.0	0	0	62.9
Sample covariance matrix (New), D	14.12	12.21	2.0	0	0	79.5
Ledoit	14.30	10.52	4.5	− 1.1	− 57.4	275.6
Ledoit, C	13.79	12.30	7.7	0	0	71.0
Ledoit, D	14.18	12.21	2.0	0	0	84.5
Fama–French 3-factor	13.07	11.25	4	− 1.4	− 66.9	276.8

Table V — *Continued*

Covariance matrix estimator	Mean	Std Dev	Max. weight	Min. weight	Short Interest	No. of Positive Weights
Fama–French 3-factor, C	13.03	12.03	12.7	0	0	40.7
Fama–French 3-factor, D	13.09	12.15	2	0	0	63.6
Panel C: Naive diversification						
Equal-weighted portfolio of the 500 stocks	14.52	17.48	0.2	0	0	500
Value-weighted portfolio of the 500 stocks	13.39	15.60	7.3	0	0	500
Equal-weighted portfolio of 25 stocks	15.16	17.78	4	0	0	25
Value-weighted portfolio of 25 stocks	14.25	17.79	30	0	0	25

mators. The out-of-sample annualized standard deviation is about 12 percentage points per year for all of the estimators, including the shrinkage estimator of Ledoit and the daily sample covariance matrix. Imposing the no-short-sales restriction leads to a small increase—between 8 and 14 percent—in the standard deviation of the minimum variance portfolios constructed using factor models. This decline in the performance is consistent with the observation by Green and Hollifield (1992) that short-sale restrictions probably do not hold in the population. The number of assets in the portfolio varies from a low of 24 for the monthly sample covariance matrix estimator to a high of 65 for the daily sample covariance matrix. In comparison, the equally weighted portfolio of the 500 stocks has an annualized standard deviation of 17 percentage points, while the value-weighted portfolio of the 500 stocks has a standard deviation of 16 percentage points. A portfolio of 25 randomly picked stocks has a standard deviation of 18 percentage points, which is 40 percent more than that of the optimal minimum variance portfolio constructed using the sample covariance matrix subject to the no-short-sale restrictions.

This clearly indicates that portfolio optimization can achieve a lower risk than naive diversification. Imposing an upper bound of two percentage points on portfolio weights in addition to a lower bound of zero does not affect the out-of-sample variance of the resulting minimum variance portfolios in any significant way.

When daily returns are used, the sample covariance matrix estimator performs the best when there are no portfolio weight constraints. The corrections for microstructure effects suggested in the literature do not lead to superior performance. When portfolio weights are constrained to be nonnegative, all the models perform about equally well.

Table VI presents the performance of the minimum tracking error portfolios. Several patterns emerge. Among the estimators that use monthly data, the

Table VI

Ex Post Mean, Standard Deviation, and Other Characteristics of the Minimum Tracking Error Variance Portfolios

At the end of April each year from 1968 to 1997, the covariance matrix of a random sample of 500 stocks is estimated according to various estimators. We use these covariance matrix estimates to construct the minimum tracking error variance portfolios, both constrained and unconstrained. The target is the S&P 500 returns. We hold the portfolios for the next 12 months and their monthly tracking errors are recorded. The ex post means and standard deviations of the tracking errors and some characteristics of the tracking portfolios are reported. The C after an estimator indicates the nonnegativity constrained portfolios, and D after an estimator indicates the portfolio with both the nonnegativity constraint and an upper bound of two percent. For the equal-weighted and value-weighted portfolios of 25 stocks, the 25 stocks are randomly selected from the 500. Means and standard deviations are in percentage per year; maximum weight, minimum weight, and short interest numbers are in percentage.

Covariance Matrix Estimator	Mean	Std Dev	Max. weight	Min. weight	Short Interest	No. of Positive Weights
Panel A: Covariance matrix estimated using monthly return data						
Sample covariance matrix, C	4.28	3.36	2.4	0	0	200
Sample covariance matrix, D	4.23	3.40	1.8	0	0	194
One-factor model	4.91	5.04	0.9	0	0	499
One-factor model, C	4.91	5.04	0.9	0	0	499
One-factor model, D	4.91	5.04	0.9	0	0	499
Ledoit	4.18	3.48	1.8	− 0.3	− 8.7	391
Ledoit, C	4.21	3.34	2.2	0	0	314
Ledoit, D	4.27	3.36	1.8	0	0	312
Fama–French 3-factor model	3.76	4.41	1.4	− 0.2	− 3.3	433.1
Fama–French 3-factor model, C	3.71	4.39	1.5	0	0	394.6
Fama–French 3-factor model, D	3.71	4.39	1.5	0	0	394.5
Panel B: Covariance matrix estimated using daily return data						
Sample covariance matrix	3.69	2.94	5.3	− 1	− 36.6	322.4
Sample covariance matrix, C	3.93	2.78	5.4	0	0	231.3
Sample covariance matrix, D	4.17	2.96	2	0	0	227.6
Sample covariance matrix of CHMSW, C	3.92	2.75	5.4	0	0	231.4
Sample covariance matrix of CHMSW, D	4.17	2.92	2	0	0	226.3
Sample covariance matrix (New)	3.72	2.92	5.3	− 1	− 36.8	322.2
Sample covariance matrix (New), C	3.93	2.73	5.4	0	0	228
Sample covariance matrix (New), D	4.17	2.89	2	0	0	224.2

Table VI — *Continued*

Covariance matrix estimator	Mean	Std Dev	Max. weight	Min. weight	Short Interest	No. of Positive Weights
Ledoit	4.20	3.31	3.7	– 0.4	– 11.1	375.0
Ledoit, C	4.31	3.30	3.8	0	0	317.4
Ledoit, D	4.39	3.41	2.0	0	0	311.1
Fama–French 3-factor	3.82	4.50	1.3	– 0.2	– 3.7	427.8
Fama–French 3-factor, C	3.76	4.50	1.5	0	0	384.1
Fama–French 3-factor, D	3.77	4.51	1.4	0	0	383.8
Panel C: Naive diversification						
Equal-weighted portfolio of the 500 stocks	4.92	6.58	0.2	0	0	500
Value-weighted portfolio of the 500 stocks	3.6	2.37	6.7	0	0	500
Equal-weighted portfolio of 25 stocks	4.66	8.62	4	0	0	25
Value-weighted portfolio of 25 stocks	4.13	8.24	29.7	0	0	25

monthly covariance matrix estimator and the Ledoit estimator perform the best. The former has only about 200 stocks, whereas the latter has about 300 stocks.

The tracking error of the one-factor model with nonnegativity constraints is 5.04 percent, which is rather large compared to 3.36 percent for the sample covariance matrix estimator. This should not be surprising, since the first dominant factor becomes less important for tracking error minimization. The portfolio weight constraints are not binding with the one-factor model. This is strong support for the Green and Hollifield conjecture that large negative (and positive) weights are due to the presence of a single dominant factor which is effectively removed when we consider tracking error minimization.

The multifactor models of Connor and Korajczyk (not reported in the table) as well as Fama and French perform better than the single-factor model. This is consistent with the observations in Chan et al. (1999). As is to be expected, portfolio weight constraints are not important when using multifactor models. However, the tracking error for the factor models is about one percentage point more than that for the sample covariance matrix and the Ledoit estimators.

With daily data, we expect the precision of all the estimators to improve. If our conjecture that portfolio weight constraints lead to better performance due to the shrinkage effect, then such constraints should become less important when daily data is used. This is what we find. When short sales are allowed, the short interest is 36.6 percent, a substantial amount. Imposing the nonnegativitiy constraint on portfolio weights reduces the number of stocks to 231. However, the performance is hardly affected. The tracking error goes down from 2.94 percent to 2.78 percent.

With daily data, the sample covariance matrix estimator dominates all factor models, with the factor models all perform equally well. Again corrections for

Table VII

T-tests of Equal Mean Returns and Equal Mean Squared Returns

This table reports the *t*-tests of equal mean returns and equal mean squared returns of the minimum variance and minimum tracking error variance portfolios. For each such portfolio, we test whether its mean returns and mean squared returns are statistically different from those of the nonnegativity constrained portfolio constructed from the sample covariance matrix of monthly returns.

	Minimum Variance Portfolio		Minimum Tracking Error Portfolio	
Covariance Matrix Estimator	Equality in Mean Return	Equality in Mean Squared Return	Equality in Mean Return	Equality in Mean Squared Return
Monthly sample covariance matrix, D	−0.01	1.36	−0.53	1.49
One-factor model	0.26	−1.16	1.06	5.35
One-factor model, C	−0.76	0.19	1.06	5.34
One-factor model, D	−0.05	0.15	1.06	5.34
Ledoit	−0.40	−2.99	−0.32	1.06
Ledoit, C	−1.05	−0.60	−0.30	−0.32
Ledoit, D	−0.07	−0.02	−0.03	0.08
Fama–French 3-factor model	−0.38	−2.34	−1.02	4.46
Fama–French 3-factor model, C	−1.08	−0.29	−1.15	4.42
Fama–French 3-factor model, D	−0.22	0.23	−1.15	4.41
Daily sample covariance matrix	0.38	−3.86	−1.30	−3.29
Daily sample covariance matrix, C	0.35	−0.11	−0.87	−4.43
Daily sample covariance matrix, D	0.65	−0.22	−0.28	−3.26
Daily sample cov matrix of CHMSW, C	0.29	−0.20	−0.89	−4.82
Daily sample cov matrix of CHMSW, D	0.62	−0.31	−0.29	−3.72
Daily sample cov matrix (New)	0.29	−3.82	−1.24	−3.40
Daily sample cov matrix (New), C	0.23	−0.35	−0.87	−5.09
Daily sample cov matrix (New), D	0.57	−0.44	−0.29	−4.11
Daily Ledoit	0.62	−4.34	0.04	−1.10
Daily Ledoit C	0.22	−0.23	0.30	−1.04
Daily Ledoit D	0.60	−0.41	0.50	−0.33
Daily Fama–French 3-factor	−0.31	−2.41	−0.82	4.28
Daily Fama–French 3-factor, C	−0.51	−0.95	−0.95	4.25
Daily Fama–French 3-factor, D	−0.44	−0.78	−0.93	4.26
Equal-weighted portfolio of the 500 stocks	0.54	6.79	0.73	6.76
Value-weighted portfolio of the 500 stocks	−0.10	5.25	−1.48	−7.33
Equal-weighted portfolio of 25 stocks	0.82	6.94	0.29	9.48
Value-weighted portfolio of 25 stocks	0.32	7.07	−0.10	9.88

microstructure effects make little difference in the ex post tracking error performance.

The value-weighted portfolio of the 500 stocks performs the best. This is to be expected since the benchmark is the S&P 500 portfolio, which is value weighted.

Table VII reports the t-tests for the difference between the mean returns and mean squared returns of the portfolios, compared with the nonnegativity-constrained portfolio from the monthly return sample covariance matrix. Since the differences in mean returns are all insignificant, the t-test for the difference in squared returns serves as a test for the difference in return variances, which is the focus of our study. We can see that for both portfolio variance minimization and tracking error minimization, once the nonnegativity constraint is imposed, the more sophisticated estimators do not, in general, give better out-of-sample performance than the monthly return sample covariance matrix. However, there is evidence that using daily returns can lead to smaller out-of-sample tracking error. Why using daily returns can lead to lower tracking error but not total risk is an issue for future investigation.

D. A Comparison of the Global Minimum Variance Portfolios and Tangency Portfolios

Earlier we justified examining the global minimum variance portfolio instead of the tangency portfolio based on results in the literature. In this section, we provide a comparison of the two portfolios that justifies our focus on the global minimum variance portfolios.

Table VIII gives the comparative results for the portfolios constructed using a random sample of 500 stocks. We only include the result for several covariance matrix estimators that use monthly return data. (The results for other estimators are qualitatively similar.) The procedure of constructing the optimal portfolios and examining their out-of-sample characteristics is the same as in the previous subsections. One difference is that here we record the ex post excess returns in excess of the one-month T-bill rate, and compare Sharpe ratios across different optimal portfolios. In terms of the out-of-sample Sharpe ratio, all the tangency portfolios perform worse than the equally weighted portfolio. The unconstrained portfolios perform uniformly worse than the constrained portfolios. For every covariance matrix estimator, the global minimum variance portfolios perform better than the tangency portfolios. The in-sample optimism (i.e., the fact that the in-sample Sharpe ratios are much higher than the out-of-sample ones) of the unconstrained portfolios is rather large, and the unconstrained tangency portfolios involve substantial short positions that would be difficult to implement in practice. For example, for the Fama–French three-factor model, the total short position is, on average, -704 percent of the portfolio's value. These problems are much less severe for the global minimum variance portfolios.

Table IX gives the result for the optimal portfolios constructed using the 25 Fama–French size/book-to-market sorted portfolios. The procedure of constructing the optimal portfolios is the same as before, except the primitive assets are the 25 portfolios. Since portfolio average returns have smaller variances com-

Table VIII

The Ex Post Performance of the Tangency and Global Minimum Variance Portfolios

At the end of April each year from 1968 to 1997, the monthly return covariance matrix of a random sample of 500 stocks is estimated according to various estimators and using returns of the past five years. The tangency and the global minimum variance portfolios are then formed and held for one year. The table reports characteristics of in- and out-of-sample excess returns of these portfolios. Means and standard deviations are those of annual excess returns in excess of the one-month T-bill rate. Average total short position is in percentage. The term "constrained" after a covariance matrix estimator means the nonnegativity constrained optimal portfolio.

	In-sample			Out-of-sample				
Portfolios	Mean	Std Dev.	Sharpe Ratio	Mean	Std Dev	Sharpe Ratio	Average Total Short Position	Average No. Of Assets Held Long
Panel A: Tangency portfolios								
Sample cov matrix, constrained	32.15	15.34	2.10	7.11	20.12	0.35	0.0	17.7
One-factor model	143.08	15.99	8.95	22.14	226.17	0.10	− 533.7	177.9
One-factor model, constrained	34.19	15.69	2.18	7.69	21.09	0.36	0.0	27.0
FF 3-factor model	185.99	20.24	9.19	− 11.97	128.16	− 0.09	− 704.3	203.3
FF 3-factor model, constrained	32.76	15.69	2.09	8.17	20.12	0.41	0.0	23.9
Ledoit estimator	264.08	28.63	9.22	3.36	107.28	0.03	− 1194.3	217.8
Ledoit estimator, constrained	33.94	16.05	2.11	7.52	20.52	0.37	0.0	24.8
Equally weighted portfolio				7.95	17.54	0.45	0.0	500.0
Panel B: Global minimum variance portfolios								
Sample cov matrix, constrained	6.83	7.06	0.97	6.83	12.43	0.55	0.0	24.0
One-factor model	2.80	2.56	1.09	7.42	11.67	0.64	− 51.0	268.0
One-factor model, constrained	5.62	6.05	0.93	6.11	12.61	0.48	0.0	39.0
FF 3-factor model	4.41	2.74	1.61	6.46	11.33	0.57	− 63.9	283.5
FF 3-factor model, constrained	6.49	6.66	0.97	6.08	12.38	0.49	0.0	39.7
Ledoit estimator	3.47	2.96	1.17	6.45	10.72	0.60	− 81.0	283.1
Ledoit estimator, constrained	6.09	6.95	0.88	6.20	12.29	0.50	0.0	39.2
Equally weighted portfolio				7.95	17.54	0.45	0.0	500.0

pared to individual stocks, they provide a more accurate estimate of the expected return on the corresponding portfolios. However, even in this case, for every covariance matrix estimator, the unconstrained global minimum variance portfolios have a higher Sharpe ratio out of sample than the corresponding unconstrained

Table IX

The Ex Post Performance of the Tangency and Global Minimum Variance Portfolios

The assets are the Fama–French 25 size/book-to-market sorted portfolios. The covariance matrix is estimated at the end of April, using monthly returns of the previous five years. The tangency portfolios and the global minimum variance portfolios are then formed and held for one year. This procedure is repeated from 1968 to 1999. The table reports characteristics of ex post excess returns of these portfolios. Means and standard deviations are those of annual excess returns in excess of the risk-free rate.

	In Sample			Out of Sample				
Portfolios	Mean	Std Dev.	Sharpe Ratio	Mean	Std Dev	Sharpe Ratio	Average Total Short Position	Average No. of Assets Held Long
Panel A: Tangency portfolios								
Sample cov matrix	61.00	17.23	3.54	28.98	82.44	0.35	− 1351.6	8.4
Sample cov matrix, constrained	13.00	16.03	0.81	8.23	16.17	0.51	0.0	3.0
One-factor model	91.32	31.29	2.92	61.77	158.63	0.39	− 784.8	10.4
One-factor model, constrained	12.90	15.60	0.83	8.10	16.32	0.50	0.0	4.0
FF 3-factor model	298.49	116.03	2.57	− 12.87	257.96	− 0.05	− 3972.4	10.2
FF 3-factor model, constrained	13.02	16.04	0.81	8.21	16.20	0.51	0.0	2.9
Ledoit estimator	211.08	97.97	2.15	39.60	253.06	0.16	− 3017.1	9.9
Ledoit estimator, constrained	13.00	15.93	0.82	8.19	16.15	0.51	0.0	3.5
Equally weighted portfolio				7.38	17.86	0.41	0.0	25.0
Panel B: Global minimum variance portfolios								
Sample cov matrix	9.06	7.44	1.22	9.63	15.93	0.60	− 481.5	13.0
Sample cov matrix, constrained	6.08	13.00	0.47	6.92	14.99	0.46	0.0	3.6
One-factor model	7.77	8.40	0.92	8.25	16.68	0.49	− 147.9	12.4
One-factor model, constrained	6.50	12.91	0.50	6.90	15.42	0.45	0.0	4.5
FF 3-factor model	8.19	8.61	0.95	10.25	14.31	0.72	− 252.4	13.0
FF 3-factor model, constrained	6.12	13.05	0.47	6.97	15.10	0.46	0.0	3.4
Ledoit estimator	7.53	9.28	0.81	8.53	13.98	0.61	− 222.6	13.4
Ledoit estimator, constrained	6.20	13.02	0.48	6.94	15.03	0.46	0.0	4.1
Equally weighted portfolio				7.38	17.86	0.41	0.0	25.0

tangency portfolios. With only 25 assets, the sample covariance matrix is nonsingular. However, imposing the constraint still does not hurt, as the constrained minimum variance portfolio has a lower variance than the unconstrained one. The constrained tangency portfolios have a higher Sharpe ratio than the corresponding global minimum variance portfolios. However, the differences are not statistically significant. This is consistent with the view that there is not enough information in the sample mean about the population mean. Hence, it is important to bring additional information about the mean for use in portfolio selection. The interested reader is referred to Black and Litterman (1991), Pastor and Stambaugh (2000), and Wang (2002) on how this can be accomplished using a Bayesian framework.

Comparing the results in the two tables, we see that for a large cross section of stocks, the global minimum variance portfolio performs better than the tangency portfolio. For the 25 Fama–French portfolios, when the nonnegativity constraint is imposed, the opposite is true, but the differences are not statistically significant. This justifies our focus on the global minimum variance portfolios when working with a large collection of stocks.

Another issue is that the constrained global minimum variance portfolios, while having a higher Sharpe ratio than the equally weighted portfolio, are likely to be less "robust," in the sense that something "bad" could happen to a few assets that affect the global minimum variance portfolios adversely but leave the equally weighted portfolio almost unaffected. This is because the constrained global minimum variance portfolios have far fewer assets—about a sixth as many as the equally weighted portfolio in the case of the 25 Fama–French portfolios and a twentieth as many assets in the 500 asset case. This issue needs to be examined in future research.

IV. Concluding Remarks

Practitioners often restrict portfolio weights to be positive when constructing minimum variance and minimum tracking error portfolios. Based on the covariance structure of U.S. stock returns, Green and Hollifield (1992) argue that minimum variance portfolios will contain extreme positive and negative weights. In that case, restricting portfolio weights to be positive as done in practice should hurt whereas the empirical evidence is to the contrary. We reconcile this apparent contradiction.

We show that constructing a minimum risk portfolio subject to the constraint that portfolio weights should be positive is equivalent to constructing it without any constraints on portfolio weights after modifying the covariance matrix in a particular way. The modification typically shrinks the larger elements of the covariance matrix towards zero. This has two effects. On the one hand, to the extent an estimated large covariance is due to sampling error, the shrinkage leads to a more precise estimate of the corresponding element in the population. On the other hand, the population covariance itself could be large, in which case the shrinkage introduces specification error. The net effect depends on the trade-off between sampling and specification errors. If the sampling error is relatively large when compared to the specification error, restricting portfolio weights to be positive can help even when the restrictions do not hold in the population. Imposing upper bounds on portfolio weights while constructing minimum risk portfolios have similar shrinkage-like effects. We examine the trade-off between sampling error and specification error using simulations.

We demonstrate through Monte Carlo simulations calibrated to U.S. stock return characteristics that the benefit to imposing nonnegativity constraints on portfolio weights can be substantial in large cross sections even when the constraints are wrong. When the no-short-sale restriction is already in place, minimum variance and minimum tracking error portfolios constructed using the sample covariance matrix perform as well as those constructed using covariance matrices estimated using factor models and shrinkage methods. The use of daily return data instead of monthly return data helps when constructing minimum tracking error portfolios but not in constructing minimum variance portfolios. Corrections for market microstructure effects that have been suggested in the literature when using daily return data do not help.

A striking feature of minimum variance portfolios constructed subject to the restriction that portfolio weights should be nonnegative is that investment is spread over only a few stocks. The minimum variance portfolio of a 500-stock universe has between 24 and 40 stocks, depending on which covariance matrix estimator was used. The annualized standard deviation of the return on the minimum variance portfolio constructed using the sample covariance matrix and 60 months of observations on returns is 12.43 percent. In contrast, the return on the equally weighted portfolio of all the 500 stocks in the universe has a standard deviation of 17.48 percent, that is, 1.4 times as large. The minimum variance portfolio also has a higher sample Sharpe ratio than the equally weighted portfolio of the 500 stocks in the universe. The return on a portfolio of 25 randomly picked stocks has a standard deviation of 17.78 percent, about the same as the equally weighted portfolio of 500 stocks in the universe. While there is little incremental benefit to naive diversification beyond 25 stocks, there is substantial benefit to picking the right 25 stocks using standard portfolio optimization methods.

Having only 25 stocks in the minimum variance portfolio should raise some concern. Variance is an adequate measure of risk of a return that has a Normal distribution. The return on a portfolio of a large collection of stocks will, in general, be close enough to Normal in distribution to justify the use of variance as the measure of risk. This may not be true for a 25-stock portfolio. In particular, the probability of an extreme event, both good as well bad, may be substantially higher than that computed by assuming a Normal distribution. Imposing upper bounds on portfolio weights would be one way to ensure that optimal portfolios will contain a sufficiently large collection of stocks.

Appendix A: Proofs

Proof of Proposition 1:[11] The matrix $\tilde{S}$ is obviously symmetric. Now we prove that it is positive semi-definite. Suppose that

$$(\omega_1, \ldots, \omega_N, \lambda_1, \ldots, \lambda_N, \delta_1, \ldots, \delta_N, \lambda_0) \equiv (\omega', \lambda', \delta', \lambda_0)$$

[11] We thank Gopal Basak for providing a part of this proof.

is a solution to the constrained portfolio variance minimization problem (1)–(4). For any vector x,

$$x'\tilde{S}x = x'Sx - x'(\mathbf{1}\lambda' + \lambda\mathbf{1}')x + x'(\mathbf{1}\delta' + \delta\mathbf{1}')x = x'Sx - 2(x'\mathbf{1})(x'(\lambda - \delta)). \quad \text{(A1)}$$

(Here $\mathbf{1}$ is a vector of ones.) By the first-order condition, $\lambda - \delta = S\omega - \lambda_0\mathbf{1}$. Hence, $x'(\lambda - \delta) = x'S\omega - \lambda_0 x'\mathbf{1}$. Therefore,

$$2(x'\mathbf{1})(x'(\lambda - \delta)) = 2(x'\mathbf{1})(x'S\omega) - 2\lambda_0(x'\mathbf{1})^2. \quad \text{(A2)}$$

But

$$|(x'\mathbf{1})(x'S\omega)| = (x'\mathbf{1})(x'S^{1/2})(S^{1/2}\omega)| \leq |(x'\mathbf{1})(x'Sx)^{1/2}(\omega'S\omega)^{1/2}.$$

The equality holds since S is positive semi-definite, and the weak inequality is due to the Cauchy–Schwarz inequality.

Again from the first order condition,

$$0 \leq \omega'S\omega = \omega'\lambda - \omega'\delta + \lambda_0\omega'\mathbf{1} = \lambda_0 - \bar{\omega}\delta'\mathbf{1} \leq \lambda_0.$$

So

$$|(x'\mathbf{1})(x'S\omega)| \leq |(x'\mathbf{1})|(x'Sx)^{1/2}(\lambda_0)^{1/2}$$

Combining the above inequality with (A1) and (A2), we have

$$\begin{aligned} x'\tilde{S}x &= x'Sx - 2(x'\mathbf{1})x'S\omega + 2\lambda_0(x'\mathbf{1})^2 \\ &\geq x'Sx - 2|(x'\mathbf{1})x'S\omega| + 2\lambda_0(x'\mathbf{1})^2 \\ &\geq x'Sx - 2|(x'\mathbf{1})|(x'Sx)^{1/2}(\lambda_0)^{1/2} + 2\lambda_0(x'\mathbf{1})^2 \\ &= (a-b)^2 + b^2, \end{aligned} \quad \text{(A3)}$$

where $a = (x'Sx)^{1/2}$ and $b = \lambda_0^{1/2}|(x'\mathbf{1})|$. Obviously this is always nonnegative. So $\tilde{S}$ is positive semi-definite.

Because $\tilde{S}$ is positive semi-definite, to show that $\omega = (\omega_1,\ldots,\omega_N)$ is a (unconstrained) global minimum variance portfolio of $\tilde{S}$, it suffices to verify the first-order condition:

$$\begin{aligned} \tilde{S}\omega &= S\omega - (\mathbf{1}\lambda' + \lambda\mathbf{1}')\omega + 2(\mathbf{1}\delta' + \delta\mathbf{1}')\omega \\ &= S\omega - \lambda\mathbf{1}'\omega + \mathbf{1}\bar{\omega}(\delta'\mathbf{1}) + \delta\mathbf{1}'\omega \\ &= S\omega - \lambda + \bar{\omega}(\delta'\mathbf{1})\mathbf{1} + \delta \\ &= (\lambda_0 + \bar{\omega}\delta'\mathbf{1})\mathbf{1}. \end{aligned}$$

The second equality follows from the fact that $\omega_i\lambda_i = 0$ for all i, and $\delta_i(\omega_i - \bar{\omega}) = 0$ for all i. The third equality holds because $\sum_j \omega_j = 1$, and the last equality follows from (5). The fact that $\tilde{S}\omega = (\lambda_0 + \bar{\omega}\delta'\mathbf{1})\mathbf{1}$ shows that ω solves the (unconstrained) variance minimization problem for covariance matrix $\tilde{S}$. Q.E.D.

Proof of Proposition 2: First, let us rewrite the second constraint in the constrained MLE problem as

$$(1-\bar{\omega})\sum_j \Omega^{i,j} \leq \bar{\omega}\sum_{k\neq i}\sum_j \Omega^{k,j}.$$

Let the constrained MLE be $\hat{\Omega}$. The Kuhn–Tucker conditions for the constrained MLE problem are (Morrison (1990))

$$\frac{\partial l}{\partial \Omega^{i,i}} = \frac{1}{2}\hat{\Omega}_{i,i} - \frac{1}{2}S_{i,i} = -\lambda_i + (1-\bar{\omega})\delta_i, \text{ all } i \tag{A4}$$

$$\frac{\partial l}{\partial \Omega^{i,j}} = \hat{\Omega}_{i,j} - S_{i,j} = -(\lambda_i + \lambda_j) + (1-\bar{\omega})(\delta_i + \delta_j), \text{ all } i<j \tag{A5}$$

$$\lambda_i \geq 0, \text{ and } \lambda_i = 0 \text{ if } \sum_j \hat{\Omega}^{i,j} > 0, \tag{A6}$$

$$\delta_i \geq 0, \; and \; \delta_i = 0 \; if \; (1-\bar{\omega})\sum_j \Omega^{i,j} < \bar{\omega}\sum_{k\neq i}\sum_j \Omega^{k,j}. \tag{A7}$$

These conditions imply that the constrained MLE can be written as

$$\hat{\Omega} = S - (\lambda\mathbf{1}' + \mathbf{1}\lambda') + (\mathbf{1}-\bar{\omega})(\delta\mathbf{1}' + \mathbf{1}\delta')$$

Notice that $\hat{\Omega}$ has the same form as $\tilde{S}$.

We will only prove Part 1. The proof of Part 2 is similar.

Let $\{\lambda_i, \delta_i, \omega_i\}_{i=1,\ldots,N}$ be a solution to the constrained portfolio variance minimization problem (1)–(4), given S, and construct $\tilde{S}$ according to (8). Then we can easily verify that $\{\lambda_i(1-\bar{\omega})\delta_i\}_{i=1,\ldots,N}$ and $\tilde{S}$ satisfies (A4), (A5), and the first halves of (A6) and (A7). Now we need to verify the second halves of (A6) and (A7), that is, $\sum_j \tilde{S}^{i,j} > 0$ implies $\lambda_i = 0$ and $(1-\bar{\omega})\sum_j \tilde{S}^{i,j} < \bar{\omega}\sum_{k\neq i}\sum_j \tilde{S}^{k,j}$ implies $\delta_i = 0$. By Proposition 1, ω is the unconstrained global minimum variance portfolio of $\tilde{S}$, so $\omega_i = \sum_j \tilde{S}^{i,j} / \sum_j \sum_k \tilde{S}^{j,k}$. From the second half of (6) we know that $\omega_i > 0$ implies $\lambda_i = 0$. This says that $\sum_j \tilde{S}^{i,j} > 0$ implies $\lambda_i = 0$. Likewise, the second half of (7) implies the second half of (A7). Q.E.D.

Appendix B: Covariance Matrix Estimators That Use Daily Returns

This appendix describes how we estimate monthly covariance matrices using daily return data after taking into account serial correlations and cross-correlations at various leads and lags induced by microstructure effects. We review the estimators proposed by Scholes and Williams (1977), Dimson (1979), and Cohen et al. (1983, hereafter CHMSW). We point out that the CHMSW estimator of the sample covariance matrix is usually not positive semi-definite. We then propose a new methodology to estimate the covariance matrices of monthly returns using daily returns that are always positive semi-definite. Our estimator uses the fact that the continuously componded monthly return is the sum of continuously compounded daily returns.

The CHMSW estimator is based on the following relationship:

$$\operatorname{cov}\left(r^t_{j,t}, r^t_{k,t}\right) = \operatorname{cov}\left(r_{j,t}, r_{k,t}\right) + \sum_{n=1}^{L} \operatorname{cov}\left(r_{j,t}, r_{k,t-n}\right) + \sum_{n=1}^{L} \operatorname{cov}\left(r_{j,t-n}, r_{k,t}\right), \tag{A8}$$

for any pair of stocks j and k, $j \neq k$. Here r^t denotes the true date t return and r denotes the observed return; t and $t-n$ denote dates t and $t-n$. Based on this relation we can estimate cov $(r^t_{j,t}, r^t_{k,t})$ using observed daily returns. In practice, L is set to three and the variances are estimated using sample variances without any adjustment (Cohen et al. (1983) and Shanken (1987)). The Scholes–Williams estimator is the special case of the above, with L set to one. The estimate of the full sample covariance matrix using the CHMSW method is denoted as "Sample Covariance Matrix of (CHMSW)" in Tables V–VII.

Equation (A8) is also valid if either asset (or both) is a portfolio. For a well-diversified portfolio, (A8) is approximately valid for its variance also. Based on this, we can estimate a stock's beta as its covariance with the market portfolio divided by the market portfolio's variance. We estimated daily one-factor models using this strategy but the results for these estimators are not reported for brevity.

There is a problem with these estimators. The estimated covariance matrix, constructed from individual covariances and variances, is not positive semi-definite. This is problematic for portfolio optimization. We propose a new estimator that does not have this problem.[12] Notice that monthly log returns are simply sums of daily log returns:

$$r_{i,r} = \sum_{t=(r-1)m+1}^{rm} r_{i,t}.$$

Here τ is month τ, t is day t, and m is the number of days in a month. Then the monthly return covariance is

$$\operatorname{cov}\left(r_{i\tau}, r_{j\tau}\right) = \sum_{t=(\tau-1)m+1}^{\tau m} \sum_{s=(\tau-1)m+1}^{\tau m} \operatorname{cov}\left(r_{i,t}, r_{j,s}\right).$$

Unlike the CHMSW estimator (A8), the above is valid for any i and j, even if either one (or both) is a portfolio. Assuming covariance stationarity as usual, we can drop the time subscripts, and get

[12] There is a difference between our approach and the approach taken by Scholes and Williams (1977) and CHMSW. While they want to estimate the "true" covariances and betas using the daily returns, we want to use daily returns to estimate the covariances and betas of monthly returns.

$$\begin{aligned}\text{cov}^M(r_i, r_j) = {} & m \cdot \text{cov}^D(r_{i,t}, r_{j,t}) \\ & + (m-1) \cdot \left(\text{cov}^D(r_i, r_{j,t+1}) + \text{cov}^D(r_{i,t+1}, r_{j,t})\right) \\ & + (m-2) \cdot \left(\text{cov}^D(r_i, r_{j,t+2}) + \text{cov}^D(r_{i,t+2}, r_{j,t})\right) \\ & + \cdots + \left(\text{cov}^D(r_{i,t}, r_{j,t+m-1}) + \text{cov}^D(r_{i,t+m-1}, r_{j,t})\right). \end{aligned} \tag{A9}$$

The superscripts M and D denote the covariance of the monthly returns and the covariance of the daily returns, respectively. We set $m = 21$, for there are about 21 trading days in one month.

An obvious approach is to use the sample counterpart of the right-hand side (RSH) of (A9) to estimate the monthly return covariance. However, to guarantee that the covariance matrix is positive semi-definite, we need to further adjust the covariance estimates (A9) slightly. Let

$$R_0 = \left(r_{t,i} - \bar{r}_{\cdot,i}\right)_{t=1,\ldots,T; i=1,\ldots,N},$$

represent the matrix of demeaned returns. For $j = 0, 1, \ldots, m-1$ let R_{-j} be the same size as R_0, with the first j rows set to zeros, and the remaining $T-j$ rows the same as R_0's first $T-j$ rows (i.e., R_{-j} is the matrix of lag-j demeaned returns). Let

$$\begin{aligned} \Omega_j &= R_0' R_{-j}/T, \\ S &= m\Omega_0 + \sum_{j=1}^{m-1} (m-j)\left(\Omega_j + \Omega_j{}'\right). \end{aligned} \tag{A10}$$

Then S is positive semi-definite (see Newey and West (1987) for the proof) and S is a consistent estimator of the RHS of (A9). The covariance matrix estimated this way is denoted as "Sample Covariance Matrix (New)" in Tables V–VII.

If we assume a k-factor model, then the beta estimates are

$$b = \left(\hat{\text{var}}(f)\right)^{-1} \hat{\text{cov}}(f, r).$$

Here b is $k \times N$, $\hat{\text{var}}(f)$ is the estimated factor covariance matrix (of size $k \times k$) and is estimated according to (A10), and $\hat{\text{cov}}(f, r)$ is estimated (similar to (A10)) by

$$\hat{\text{cov}}(f, r) = \frac{1}{T}\left[mF_0' R_0 + \sum_{j=1}^{m-1} (m-j)\left(F_0' R_{-j} + F_{-j}' R_0\right)\right],$$

with F_{-j} defined similarly as R_{-j}.

For factor models, the residual covariance matrix is assumed to be diagonal and the residual variances are estimated by the sample variances of residuals calculated from observed stock returns, observed factor returns, and the estimated betas. In Tables V–VII, the estimators denoted as "Daily Fama–French 3-Factor Model" are estimated using this strategy. We also examined the performance of a daily one-factor model, a daily Connor–Korajczyk three-factor model, and a daily Connor–Korajczyk five-factor model, all estimated using this strategy. These results are not reported for brevity.

We construct the daily Connor–Korajczyk factors and Fama–French factors by following the same procedure as outlined in Connor and Korajczyk (1988) (using daily returns instead of monthly returns) and Fama and French (1993).

REFERENCES

Best, Michael J., and Robert R. Grauer, 1991, On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results, *Review of Financial Studies* 4, 315–342.

Black, Fisher, and Robert Litterman, 1992, Global portfolio optimization, *Financial Analysts Journal* 48, 28–43.

Bloomfield, Ted, Richard Leftwich, and John Long, 1977, Portfolio strategies and performance, *Journal of Financial Economics* 5, 201–218.

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.

Cohen, Kalman, Gabriel Hawawini, Steven Maier, Robert Schwartz, and David Whitcome, 1983, Friction in the trading process and the estimation of systematic risk, *Journal of Financial Economics* 12, 263–278.

Connor, Gregory, and Robert A. Korajczyk, 1986, Performance measurement with the arbitrage pricing theory, *Journal of Financial Economics* 15, 373–394.

Connor, Gregory, and Robert A. Korajczyk, 1988, Risk and return in an equilibrium APT: Applications of a new test methodology, *Journal of Financial Economics* 21, 255–289.

Dimson, Elroy, 1979, Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7, 197–226.

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.

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

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

Green, Richard C., and Burton Hollifield, 1992, When will mean-variance efficient portfolios be well diversified? *Journal of Finance* 47, 1785–1809.

Jagannathan, Ravi, and Tongshu Ma, 2002, Risk reduction in large portfolios: Why imposing the wrong constraints helps, NBER working paper 8922.

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

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

Jorion, Philippe, 1985, International portfolio diversification with estimation risk, *Journal of Business* 58, 259–278.

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

Jorion, Philippe, 1991, Bayesian and CAPM estimators of the means: Implications for portfolio selection, *Journal of Banking and Finance* 15, 717–727.

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, Ludan, 2001, Portfolio constraints and shrinkage estimation, Working paper, Carroll School of Management, Boston College.

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

Markowitz, H. M., 1959, *Portfolio Selection: Efficient Diversification of Investments* (Yale University Press, New Haven, CT).

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

Morrison, Donald F., 1990, *Multivariate Statistical Methods* (McGraw-Hill, New York).

Newy, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Pastor, Lubos, and Robert Stambaugh, 2000, Comparing asset pricing models: An investments perspective, *Journal of Financial Economics* 56, 335–381.

Scholes, Myron, and Joseph Williams, 1977, Estimating betas from nonsynchronous data, *Journal of Financial Economics* 5, 309–328.

Schwert, G. Williams, and Paul L. Seguin, 1990, Heteroskedasticity in stock returns, *Journal of Finance* 45, 1129–1155.

Shanken, Jay, 1987, Nonsynchronous data and the covariance-factor structure of returns, *Journal of Finance* 42, 221–231.

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

Wang, Zhenyu, 2002, A shrinkage approach to model uncertainty and asset allocation, Working paper, Columbia University.