# Shrinking the Covariance Matrix

*Simpler is better.*

David J. Disatnik and Simon Benninga

The computational aspects of finding efficient portfolios have been a concern of the finance profession since the seminal work of Markowitz [1952, 1959]. While the mathematics of efficient portfolios are relatively simple, the traditional practical implementation of the theory often leads to questionable results—in many cases the covariance matrix is not invertible, and in other cases the "optimal" portfolios have very large short sale positions.

Essentially there are two approaches to deal with the problematic implementation results. The first is the theoretical approach, where one reexamines theoretical aspects and assumptions regarding portfolio optimization. The second is the implementation approach, which largely rests on the fact that the two main elements of Markowitz's mean-variance (MV) theory—the expected stock returns vector and the covariance matrix of the stock returns—are unknown, and thus must be estimated. Our work is in this implementationally oriented literature, and it deals mainly with estimation of the covariance matrix.

Estimation of the covariance matrix, like any other estimation process, is subject to error, whether sampling error or specification error. Sampling error occurs when there are not enough degrees of freedom per estimated parameter, or in other words when there are not enough observations in the sample compared to the number of estimated parameters. Specification error occurs when some form of structure is imposed on the model used in the estimation process, so the estimator becomes too specific in comparison with reality.

The traditional and probably most intuitive estimator of the covariance matrix is the sample covariance

**DAVID J. DISATNIK**
is a Ph.D. student in finance in the Faculty of Management at Tel Aviv University in Israel.
**daveydis@post.tau.ac.il**

**SIMON BENNINGA**
is a professor of finance in the Faculty of Management at Tel Aviv University in Israel.
**benninga@post.tau.ac.il**

(henceforth the sample matrix). Pafka, Potters, and Kondor [2004] state that this estimator often suffers from the *curse of dimensions*; in many cases the stock return time series period used for estimation ($T$) is not long enough compared to the number of stocks considered ($N$). As a result, the estimated covariance matrix is ill conditioned. Michaud [1989] points out that inverting such a matrix (as required by MV theory) amplifies the sampling error tremendously. Furthermore, when $N$ is bigger than $T$, the sample covariance matrix is not invertible at all.[1]

While the literature dealing with ways to improve the estimation of the covariance matrix is too extensive to survey here, several authors recently have concentrated on estimators using monthly data, assuming both return stationarity and that sample variances are good estimators of the stock variances. A fundamental principle of statistical theory is that there is a trade-off between sampling error and specification error. Hence, in order to develop a better estimator, one must reduce the huge sampling error of the sample matrix without creating too much specification error.

The findings of Chan, Karceski, and Lakonishok [1999], Bengtsson and Holst [2002], Jagannathan and Ma [2003], Ledoit and Wolf [2003, 2004], and Wolf [2004] suggest that the best estimators of this type are shrinkage estimators and portfolios of estimators.

The roots of the shrinkage method in statistics are not related to covariance estimation and can be found in the seminal work of Stein [1955].[2] Roughly speaking, in our context, a shrinkage estimator is usually a weighted average of the sample matrix with an invertible covariance matrix estimator on which quite a lot of structure is imposed and whose diagonal elements are the sample variances.[3]

The proportions used in the shrinkage estimator are often found by minimizing the quadratic risk (of error) function of the combined estimator. These proportions are supposed to guarantee a reduction in the sampling error of the sample matrix without creating instead too much specification error (which is related to the second estimator in the weighted average). As a result, the off-diagonal elements of the shrinkage estimator are moderated (or shrunk) compared to the typically large off-diagonal elements of the sample matrix. The variance elements in the diagonal are kept untouched.

Jagannathan and Ma [2000] use the concept of a portfolio of covariance estimators. A portfolio of estimators is an estimator consisting of an equally weighted average of the sample matrix and several other estimators of the covariance matrix whose diagonal elements are the sample variances and where at least one of them is invertible. This concept has also been adopted by Bengtsson and Holst [2002]; it depends on the logic that estimators based on different assumptions make errors in different directions. The portfolio of estimators diversifies among these errors, and it is hoped they cancel out.

In essence, the portfolio approach builds on the trade-off between the sampling and the specification errors. By averaging the sample matrix (which suffers from considerable sampling error) with other estimators whose primary error is specification error, one can obtain an improved covariance matrix.

Both the shrinkage estimators and the portfolios of estimators that appear in the literature have been shown to perform substantially better than the sample matrix, but the shrinkage estimators are more complex, at least in their theoretical derivation. While a portfolio of estimators is often derived simply by using an equally weighted average, the derivation of a shrinkage estimator involves solving a minimum problem for finding the proportions in the weighted average and estimating these proportions, as they depend on some unknown parameters.[4]

We examine whether there is any gain in terms of performance from using the more sophisticated shrinkage methods. We do this by running a performance contest, or a horse race between several shrinkage estimators and portfolios of estimators. We use the ex post standard deviation of the global minimum-variance portfolio (GMVP) as our criterion of improvement.

We show that all the estimators perform within the same range, and that it is actually impossible to claim that one of them is better than the other. Our conclusion is that one can use the simpler estimators rather than the more complicated ones. That is, there is no real need to use the shrinkage estimators; one can simply use the portfolios of estimators.

A significant drawback of many covariance matrix estimators, including the shrinkage estimators and the portfolios of estimators, is that they generate minimum-variance portfolios incorporating significant short sale positions. Short-selling is a significant implemental problem in portfolio computations; it is widely prohibited (mutual funds, for example, in many cases are not allowed to short sell), and many individual investors find short-selling onerous or impossible. Therefore, to the extent that short sales are considered an undesirable feature of portfolio optimization, there is some interest in finding an estimator of the covariance matrix that performs substantially better than the sample matrix and produces positive efficient portfolios.

Probably the most intuitive way to obtain positive portfolios is to add short sale constraints to the portfolio selection problem.[5] To see whether this way can also produce estimators that generate relatively low ex post standard deviations of the GMVP, we run a new contest—this time imposing the short sale constraints, and using the estimators: the sample matrix, one shrinkage estimator, one portfolio of estimators, and a matrix with only the diagonal elements of the sample matrix (henceforth the diagonal matrix), which serves as our "stalking horse," as it is subject to much specification error.[6]

Much like Bengtsson and Holst [2002] and Jagannathan and Ma [2003], we find that when the short sale constraints are imposed, all three estimators perform substantially better than the diagonal matrix. Not surprisingly, however, imposing short sale constraints has a cost. When compared to the unconstrained GMVP constructed from the shrinkage estimators and the portfolios of estimators, the GMVP has significantly higher standard deviation in the presence of the short sale constraints. This is true no matter which of the three estimators we use. This statistically significant gap between the performances of the estimators when the short sale constraints are imposed and not imposed is the cost of not holding short sale positions.[7]

Our findings differ from those of Bengtsson and Holst [2002] and Jagannathan and Ma [2003] in at least one significant respect. When the short sale constraints are imposed, both the shrinkage estimator and the portfolio of estimators perform statistically significantly better than the sample matrix. We also find that the portfolio of estimators performs at least as well as the more sophisticated shrinkage estimator. This again confirms our notion that simpler is better, at least when it comes to shrinkage.

## DATA AND METHODOLOGY

We use monthly returns on stocks traded on the New York Stock Exchange (NYSE). The stock returns are extracted from the Center for Research in Security Prices (CRSP) database. The study runs from January 1964 through December 2003.

For evaluating the performance of the estimators included in our performance contest, one can compare the estimators computed over a certain sample period (the in-sample period) with the covariance matrix realized over a subsequent period (the out-of-sample period). Our main interest is in assessing how the performance of the estimators translates into the performance of the optimal portfolios obtained from the MV optimization process. Therefore, we find it more useful to conduct an empirical performance contest that focuses on the ex post performance of the respective optimal portfolios.

We follow Chan, Karceski, and Lakonishok [1999], Bengtsson and Holst [2002], Jagannathan and Ma [2003], and Ledoit and Wolf [2003], and use the ex post standard deviation of the GMVP as our criterion of improvement. We benefit from the fact that finding the GMVP does not require estimation of the expected stock returns vector, which is beyond the scope of our research.[8]

We mimic an investor who invests in the GMVP of stocks traded on the NYSE. Our investor chooses a covariance matrix estimator that is based on historical data of stock returns and then chooses the length of the in-sample period, the historical time frame over which monthly return data are collected. The investor uses these data to compute the covariance matrix estimator, finds the GMVP, and invests in this portfolio.

Our investor has a fixed investment horizon during which the portfolio is kept unchanged (the out-of-sample period). When this period is over, the portfolio is liquidated, and the whole process begins again: estimating the covariance matrix, constructing the GMVP, and holding it until liquidation once again.

To illustrate, let us assume that the first time our investor wishes to invest is January 1974; the in-sample period is 120 months, and the out-of-sample period is 12 months. Then:

1. We collect monthly return data of stocks traded on the NYSE for 1964-1973. We choose an estimator of the covariance matrix, which is computed on the basis of these data.
2. We construct the GMVP from the estimator computed in phase 1.
3. We record the monthly returns on the GMVP for the 12 months of 1974.
4. We start the whole process all over again. We collect monthly return data of stocks traded on the NYSE for 1965-1974, and on the basis of these data we compute the same estimator used in phase 1, construct the GMVP, record its monthly returns for the 12 months of 1975, and so on.
5. We repeat the process of computing the covariance matrix, constructing the GMVP, and recording its monthly returns in the out-of-sample period 30 times (the last monthly return recorded is for December 2003). As a result, altogether, we collect 360 monthly returns on the GMVP (1974-2003).

6. We compute the standard deviation of the collected 360 monthly returns. This ex post standard deviation represents the risk our investor was exposed to in the 30 years the investment strategy was run. Given the chosen in-sample and out-of-sample periods, we can refer to the computed ex post standard deviation as a proxy of the performance of the specific estimator used for estimating the covariance matrix.

We conduct our test for several estimators. As our investor wants to minimize the risk of the investment, the lower the standard deviation of the collected 360 monthly returns, the better the estimator of the covariance matrix.[9]

We run this horse race six times, each time changing the length of the in-sample period or the length of the out-of-sample period. We use in-sample periods of 120 months and 60 months.[10] We use out-of-sample periods of 12 months, 24 months, and 36 months. We choose these three out-of-sample periods to correspond to realistic investment horizons.

As an aside, we always construct the first GMVP in January 1974, and record the last return data on the last GMVP in December 2003. Thus, in each of the six runs for every one of the seven estimators participating in our contest, we have a set of 360 monthly returns used to compute the respective ex post standard deviation.

It is worth mentioning that each time we construct a GMVP, we construct it only with NYSE stocks whose returns cover the entire in-sample and out-of-sample periods. For example, in the case of the in-sample period of 120 months and the out-of sample period of 12 months, to construct the GMVP of January 1974, we use only NYSE stocks with monthly return data for all the 132 months from January 1964 through December 1974. To construct the GMVP of January 1975, we use only NYSE stocks with monthly return data for all the 132 months from January 1965 through December 1975, and so on. Therefore, the number of stocks used for constructing the GMVP varies across the years and the runs of the contest.[11]

## COVARIANCE MATRIX ESTIMATORS

We use seven covariance matrix estimators.

***Shrinkage to the single-index model.*** Shrinkage to the single-index is the shrinkage estimator suggested by Ledoit and Wolf [2003], where the covariance matrix estimator obtained from Sharpe's [1963] single-index model (henceforth the single-index matrix) joins the sample matrix in the weighted average.[12] This estimator performed best in Ledoit and Wolf's [2003] contest and Jagannathan and Ma's [2003] contest.[13]

***Shrinkage to the constant-correlation model.*** Shrinkage to the constant-correlation is the shrinkage estimator suggested by Ledoit and Wolf [2004], where the covariance matrix estimator obtained by assuming that each pair of stocks has the same correlation (henceforth the constant-correlation matrix) joins the sample matrix in the weighted average. Ledoit and Wolf [2004] find the performance of this estimator comparable to the performance of the shrinkage to the single-index model estimator.

***A portfolio of the sample matrix, the single-index matrix, and the diagonal matrix.*** This portfolio estimator was introduced by Jagannathan and Ma [2000] and then adopted by Bengtsson and Holst [2002]. It consists of an equally weighted average of the sample matrix, the single-index matrix, and the diagonal matrix. Its performance was found to be one of the best in Bengtsson and Holst's [2002] contest.

***A portfolio of the sample matrix, the single-index matrix, and the constant-correlation matrix.*** This is an estimator so far not used in the literature. It consists of an equally weighted average of the sample matrix, the single-index matrix, and the constant-correlation matrix.

***A portfolio of the sample matrix, the single-index matrix, the constant-correlation matrix, and the diagonal matrix.*** This too is a new estimator. It consists of an equally weighted average of the sample matrix, the single-index matrix, the constant-correlation matrix, and the diagonal matrix.

***A random average of the sample matrix and the single-index matrix.*** Like the shrinkage to the single-index model estimator, this random average estimator is based on a weighted average of the sample matrix and the single-index matrix. This time the proportion of the single-index matrix in the weighted average is drawn from a uniform distribution on the interval (0.5, 1.0). We include this estimator in the contest because it helps us highlight the problem of estimating the proportions of the estimators in the shrinkage estimators. Ledoit and Wolf [2003] and Bengtsson and Holst [2002] observe that there is more sampling error in the sample matrix than there is specification error in the single-index matrix. Therefore, we allow the proportion of the single-index matrix in the random average to obtain only values greater than 0.5.

***The diagonal matrix.*** The diagonal matrix estimator has considerable specification error, and clearly does not

obey the statistical principle of reducing the sampling error of the sample matrix without creating instead too much specification error. It is included in the contest as our *stalking horse*, because for our data the sample matrix is not invertible and therefore cannot be used.

## PERFORMANCE CONTEST WITH NO SHORT SALE CONSTRAINTS

In the results of our horse race, we show empirically that there is no statistically significant gain from using the more sophisticated shrinkage methods. All the methods discussed here lead to similar improvements in terms of the ex post standard deviation of the GMVP. We conclude that simpler is better, at least when it comes to shrinkage.

Exhibit 1 reports the ex post standard deviations obtained for each covariance matrix estimator when short sale constraints are not imposed. The standard deviations are annualized through multiplication by $\sqrt{12}$.

All six estimators perform substantially better than the stalking horse, the diagonal matrix. The most important point of the comparison, however, is not the improvement of the six methods over the obviously inferior diagonal matrix, but that all the estimators we examine perform within the same range.

In other words, we find very little qualitative difference among the estimators. The widest gap in a specific run between the standard deviations corresponding to the best and worst improvements of our six estimators is obtained in the case of in-sample period of 120 months and out-of-sample period of 36 months, and it is only 0.19%. Checking, for each run, whether the tiny differences in the performance of the various estimators are statistically significant results in a negative answer in all cases.[14]

In addition, we can see that the ranking of the estimators changes from one run to another. The portfolio of the sample matrix, the single-index matrix, the constant-correlation matrix, and the diagonal matrix wins the contest four times, while the shrinkage to the single-index model estimator wins only twice.

We conclude that all six estimators produce substantially the same performance improvement, and that therefore there is no need to use the methodologically complex shrinkage estimators. Instead, we can simply use a portfolio of estimators, as long as the portfolio obeys the statistical principle of reducing the sampling error of the sample matrix without creating too much specification error.

One could claim that if a better estimator than the single-index matrix or the constant-correlation matrix is found, the shrinkage estimator relaying on this estimator will perform better than any portfolio of estimators. We do not agree with this idea. In our opinion, placing this estimator in a portfolio of estimators will result in performance within the same range as the shrinkage estimator.

An example can be found in Bengtsson and Holst [2002]. They develop a rather complicated shrinkage estimator that generates the estimator joining the sample matrix from principal components analysis. According to them, their new shrinkage estimator performs better than the

## EXHIBIT 1

**Ex Post Standard Deviations—No Short Sales Constraints**

| | In sample 120 months | | | In sample 60 months | | |
|---|---|---|---|---|---|---|
| | Out of sample period | | | | | |
| | 12 months | 24 months | 36 months | 12 months | 24 months | 36 months |
| Average size of universe of stocks | 901 | 853 | 806 | 1,261 | 1,186 | 1,110 |
| Largest universe | 1,063 | 945 | 865 | 1,739 | 1,572 | 1,424 |
| Smallest universe | 795 | 769 | 729 | 1,029 | 985 | 958 |
| | | | | | | |
| Diagonal | 13.12% | 13.12% | 13.16% | 12.90% | 12.92% | 12.84% |
| Shrinkage to constant correlation | 8.52% | 8.97% | 8.91% | 8.46% | 8.85% | 8.83% |
| Random average of sample and single index | 8.51% | 8.93% | 9.00% | 8.47% | 8.83% | 8.89% |
| Portfolio of sample, single index, constant correlation | 8.47% | 8.90% | 8.85% | 8.40% | 8.83% | 8.81% |
| Portfolio of sample, single index, constant correlation, diagonal | 8.46% | 8.85% | 8.81% | 8.37% | 8.81% | 8.78% |
| Portfolio of sample, single index, diagonal | 8.39% | 8.89% | 8.93% | 8.34% | 8.94% | 8.91% |
| Shrinkage to single index | 8.37% | 8.89% | 8.94% | 8.31% | 8.91% | 8.90% |
| Gap between worst and best improver | 0.15% | 0.12% | 0.19% | 0.16% | 0.13% | 0.13% |

shrinkage estimator of Ledoit and Wolf [2003]. Yet when they use a portfolio of estimators consisting of the estimator generated from principal components analysis, the sample matrix, and the diagonal matrix, they obtain an estimator that performs at least as well as their shrinkage estimator.

We can see that both the random average of the sample matrix and the single-index matrix and the shrinkage to the single-index model estimator perform within the same range. Theoretically, the shrinkage estimator should perform better than any other weighted average of the two estimators, as the proportions in the weighted average of the shrinkage estimator are obtained from minimizing the quadratic risk (of error) function of the combined estimator.

Yet it seems that, in practice, estimating these specific proportions gives rise to a new type of error, and overall the shrinkage estimator does not perform better than the random average. This result is similar to a result in Jagannathan and Ma [2003]. We suspect Bengtsson and Holst [2002] report a contradictory result, because they use for their random average a uniform distribution on the interval (0, 1) and not on the interval (0.5, 0.1). Thus, they do not prevent cases where the proportion of the single-index matrix is lower than the proportion of the sample matrix. In these cases, the estimator obtained has probably too much sampling error, and therefore cannot compete with the shrinkage estimator.

As an aside, we can see that when we keep the length of the out-of-sample period fixed and change the in-sample period from 120 to 60 months, we obtain quite similar results for the performance of the various estimators. When we keep the in-sample period fixed and change the out-of-sample period from 12 months to 24 or 36 months, the annualized standard deviations grow by approximately 0.45 percentage points, and the p-values for the chi-square tests for significance of differences in performance range from 2.02% to 12.71%.

Further research might address more carefully the effects (if any) of the length of the in-sample and out-of-sample periods on the performance of the covariance matrix estimators.

## PERFORMANCE CONTEST WITH SHORT SALE CONSTRAINTS

So far we have evaluated the covariance matrix estimators without addressing the issue of short-selling positions. Here we discuss the effects of imposing short sale constraints on the portfolio selection problem.

### Average Short Positions

Exhibit 2 presents the average short positions obtained for each estimator in all six runs of our contest. The amount of short positions is defined as the sum of all negative portfolio weights.

We can see that, except for the estimator based on the diagonal matrix, which generates a positive GMVP, all other estimators in all the runs of the contest generate portfolios with significant short sale positions. Moving from an in-sample period of 120 months to an in-sample period of 60 months reduces the average short positions for all estimators, but even then there are still quite significant short sale positions. Bengtsson and Holst [2002], Jagannathan and Ma [2003], and Ledoit and Wolf [2003] also report significant short positions in their performance contests.

## EXHIBIT 2
**Average Short Positions**

| *Covariance Matrix Estimator* | *In Sample/Out of Sample* | | | | | |
|---|---|---|---|---|---|---|
| | 120 / 12 | 120 / 24 | 120 / 36 | 60 / 12 | 60 / 24 | 60 / 36 |
| Diagonal | 0% | 0% | 0% | 0% | 0% | 0% |
| Random average of sample and single index | -90% | -91% | -88% | -62% | -63% | -64% |
| Portfolio of sample, single index, constant correlation | -117% | -117% | -117% | -91% | -92% | -94% |
| Portfolio of sample, single index, diagonal | -90% | -90% | -90% | -65% | -66% | -68% |
| Portfolio of sample, single index, constant correlation, diagonal | -101% | -101% | -100% | -84% | -84% | -86% |
| Shrinkage to constant correlation | -129% | -130% | -131% | -93% | -93% | -96% |
| Shrinkage to single index | -97% | -98% | -98% | -67% | -67% | -69% |

*An average short interest of −97%, obtained for the shrinkage to the single-index model estimator in the run of in-sample period of 120 months and out-of-sample period of 12 months, means that on average, over the 30 portfolios constructed in this run based on this estimator, for every dollar invested in the portfolio we short 97 cents worth of stocks, while buying $1.97 worth of other stocks.*

To the extent that short sales are considered an undesirable feature of portfolio optimization, the shrinkage estimators and the portfolios of estimators cannot satisfy us any longer. We cannot count on the diagonal estimator either, because it generates relatively high ex post standard deviations (see Exhibit 1).

Hence, our goal now is to find an estimation method that generates both low ex post standard deviations and positive portfolios.

### Imposing Short Sale Constraints

We now run our performance contest again, this time imposing the short sale constraints. We focus this time on three covariance matrix estimators—the sample matrix, the shrinkage to the single-index model estimator, and the portfolio of the sample matrix, the single-index matrix, the constant-correlation matrix, and the diagonal matrix.[15] As before, our stalking horse is the diagonal matrix, and the ex post standard deviation of the GMVP is used as our improvement criterion.

Exhibit 3 reports the ex post standard deviations obtained for each covariance matrix estimator when the short sale constraints are imposed. The standard deviations are annualized through multiplication by $\sqrt{12}$.

As in Bengtsson and Holst [2002] and Jagannathan and Ma [2003], our results confirm that imposing the short sale constraints substantially reduces the ex post standard deviations compared to the ex post standard deviation obtained for the diagonal matrix. Not surprisingly, however, imposing short sale constraints has a cost, which is seen when we compare the ex post standard deviations generated by the shrinkage estimator and the portfolio of estimators with and without short sale constraints. In the case of the shrinkage estimator, imposing the constraints increases the ex post standard deviations by about 1.3 to 1.6 percentage points. In the case of the portfolio of estimators, imposing the constraints increases the ex post standard deviations by about 1.2 percentage points for in-sample periods of 120 months and by about 0.8 percentage points for in-sample periods of 60 months.

All gaps in all runs for both estimators are statistically significant. This statistically significant gap between the performance of the estimators when the short sale constraints are imposed and not imposed is the cost of not holding short sale positions.[16]

Our findings differ from those of Bengtsson and Holst [2002] and Jagannathan and Ma [2003] in at least one significant respect. When the short sale constraints are imposed, both the shrinkage estimator and the portfolio of estimators perform statistically significantly better than the sample matrix.[17] Applying the same statistical significance tests to the results and samples of Bengtsson and Holst [2002] and Jagannathan and Ma [2003], however, reveals that in these studies the gaps obtained are not statistically significant. We believe future research should address this issue more carefully.

We can also see that the portfolio of estimators systematically generates lower ex post standard deviations than the shrinkage estimator, with p-values that range from 4.18% to 31.00%. This again confirms our notion that simpler is better, at least when it comes to shrinkage.

As an aside, note that when we keep the out-of-sample period unchanged and reduce the in-sample period from 120 months to 60 months, the portfolio of estimators performs better; the performance of the shrinkage to the single-index model estimator is almost unchanged, and the sample matrix deteriorates. This time the lowest

## EXHIBIT 3
**Ex Post Standard Deviations—Short Sales Constraints**

| | In sample 120 months | | | In sample 60 months | | |
|---|---|---|---|---|---|---|
| | Out of sample period | | | | | |
| | 12 months | 24 months | 36 months | 12 months | 24 months | 36 months |
| Average size of universe of stocks | 901 | 853 | 806 | 1,261 | 1,186 | 1,110 |
| The largest universe | 1,063 | 945 | 865 | 1,739 | 1,572 | 1,424 |
| The smallest universe | 795 | 769 | 729 | 1,029 | 985 | 958 |
| | | | | | | |
| Diagonal | 13.12% | 13.12% | 13.16% | 12.90% | 12.92% | 12.84% |
| Sample | 10.74% | 10.87% | 11.08% | 10.84% | 11.23% | 11.48% |
| Shrinkage to single index | 9.94% | 10.19% | 10.23% | 9.75% | 10.24% | 10.17% |
| Portfolio of sample, single index, constant correlation, diagonal | 9.65% | 10.01% | 10.06% | 9.15% | 9.65% | 9.57% |

p-value obtained is 6.12%. In addition, keeping the in-sample period unchanged and increasing the out-of-sample period from 12 months to 24 or 36 months damages the performance of the three estimators in terms of the ex post standard deviations. The lowest p-value obtained for this set of checks is also 6.12%.

## SUMMARY

We examine estimation of the covariance matrix of stock returns, one of the two main elements of the mean-variance theory of Markowitz [1952, 1959] for portfolio selection.

Estimating the covariance matrix solely on the basis of the sample matrix is famously difficult, as very often the sample matrix suffers from the "curse of dimensions." This has prompted a significant finance literature, looking for better ways to estimate the covariance matrix. Out of this rich literature, we have chosen to focus on estimators using monthly data and assuming both return stationarity and that sample variances are good estimators of stock variances. Combining the findings of recent studies reveals that the best estimators of that type are the shrinkage estimators and the portfolios of estimators.

In our horse race between various shrinkage estimators and portfolios of estimators, we use the ex post standard deviation of the global minimum-variance portfolio (GMVP) as our criterion of improvement. We show empirically that all the estimators perform within the same range, and that it is impossible to claim that one of them is better than the other. Hence, there is no statistically significant gain from using the more sophisticated shrinkage methods, so one can instead use the simpler portfolios of estimators. We conclude that simpler is better when it comes to shrinkage.

A significant drawback of many covariance matrix estimators, including the shrinkage estimators and the portfolios of estimators, is that they generate minimum-variance portfolios incorporating significant short sale positions. To the extent that short sales are considered an undesirable feature of portfolio optimization, the most logical way to overcome them is to add to the portfolio selection problem short sale constraints that prevent the portfolio weights from being negative, no matter which covariance matrix estimator is used.

We examine the imposition of short sale constraints, where again the ex post standard deviation of the GMVP is used as the improvement criterion. Our findings in this case differ from results elsewhere in at least one significant respect. When we impose short sale constraints, both the shrinkage estimator and the portfolio of estimators perform statistically significantly better than the sample matrix. In addition, the portfolio of estimators performs, ex post, at least as well as the more sophisticated shrinkage estimator. This again confirms our notion that simpler is better when it comes to shrinking the covariance matrix.

## ENDNOTES

The authors thank Dessi Pachamanova, Michael Wolf, and seminar participants at the University of Basel, the Hebrew University of Jerusalem, the University of Konstanz, and Tel Aviv University for helpful comments. The authors also thank the Kurt Lion Foundation for research support.

[1]See, for example, Ledoit and Wolf [2003].

[2]Efron and Morris [1977] give a beautiful description of Stein's work.

[3]It is straightforward to show that a weighted average of two matrices, one of which is invertible, is also invertible. Thus, the shrinkage estimator is always invertible.

[4]An example of the proportions estimator can be found in Ledoit and Wolf [2003]. While the shrinkage estimators may appear to be computationally complex, these are easily computed using Matlab.

[5]Jagannathan and Ma [2003] show that imposing such constraints can be interpreted as a means of shrinking. They show that constructing the constrained GMVP from the sample matrix is equivalent to constructing the unconstrained GMVP from a shrunk covariance matrix.

[6]Note that when the short sale constraints are imposed, the numerical solution for the weights of the GMVP does not depend on inverting the covariance matrix. In this case, therefore, the sample matrix can be added to the performance contest whatever the number of stocks considered.

[7]Levy and Ritov [2001] also discuss the cost of not holding short positions. They measure this cost by comparing the Sharpe ratios of the optimal portfolios when the short sale constraints are imposed and not imposed.

[8]Basak, Jagannathan, and Ma [2004] state that the estimate of the variance of a GMVP constructed using an estimated covariance matrix will on average be strictly lower than its true variance. Elton, Gruber, and Spitzer [2005] point out that evaluation of the performance of the covariance matrix estimators based on the ex post GMVP suffers from some bias, because the stocks that enter the GMVP are principally the ones with low correlations.

[9]It might appear more intuitive to calculate the ex post covariance matrix, then find the ex post GMVP generated from the ex post covariance matrix, and finally check which covariance matrix estimator generates the GMVP closest to the ex post GMVP. We cannot conduct this type of test, however,

when there are more stocks than the length of the stock return time series (as in our dataset), because then the ex post covariance matrix is singular—which prevents finding the ex post GMVP.

[10]Jobson and Korkie [1981] mention rules of thumb regarding lengths of in-sample periods of four to seven years and eight to ten years.

[11]We are aware that this common procedure introduces survivorship bias into the estimation procedure. Given that survivorship bias is common to all the compared estimators, though, we do not consider this a significant problem.

[12]We use the value-weighted portfolio of stocks included in our study as the index.

[13]In fact, in Jagannathan and Ma's [2003] contest, the shrinkage estimator performed best together with the sample covariance matrix based on daily return data, which is beyond the scope of our study.

[14]We use chi square tests throughout for statistical inference. The lowest p-value obtained is 29.12%.

[15]To find the GMVP weights, we use the Mosek optimization software together with Matlab. When the diagonal matrix is used, the short sale constraints are of course not needed, as the diagonal matrix generates a positive GMVP.

[16]The highest p-value obtained this time is 0.39%.

[17]In the in-sample period of 120 months and out-of-sample period of 24 months, for the gap between the shrinkage estimator and the sample matrix we obtain a p-value of 2.87%. In all the other cases, the highest p-value obtained is 1.07%.

## REFERENCES


Basak, G. K., R. Jagannathan, and T. Ma. "Assessing the Risk in Sample Minimum Risk Portfolios." Working paper, University of Bristol, Northwestern University and University of Utah, 2004.

Bengtsson, C., and J. Holst. "On Portfolio Selection: Improved Covariance Matrix Estimation for Swedish Asset Returns." Working paper, Lund University and Lund Institute of Technology, 2002.

Chan, L. K. C., J. Karceski, and J. Lakonishok. "On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model." *Review of Financial Studies*, vol. 12 (1999), pp. 937-974.

Efron, B., and C. Morris. "Stein's Paradox in Statistics." *Scientific American*, vol. 236 (1977), pp. 119-128.

Elton, E. J., M. J. Gruber, and J. Spitzer. "Improved Estimates of Correlation Coefficients and their Impact on the Optimum Portfolios." Working paper, New York University, 2005.

Jaggannathan, R., and T. Ma. "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps." *Journal of Finance*, vol. 54, no. 4 (2003), pp. 1651-1683.

———. "Three Methods for Improving the Precision in Covariance Matrix Estimation." Working paper, 2000.

Jobson, J.D., and B. Korkie. "Putting Markowitz Theory to Work." *The Journal of Portfolio Management*, vol. 7 (1981), pp. 70-74.

Ledoit, O., and M. Wolf. "Honey, I Shrunk the Sample Covariance Matrix." *The Journal of Portfolio Management*, vol. 30, no. 4 (2004), pp. 110-119.

———. "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection." *Journal of Empirical Finance*, vol. 10 (2003), pp. 603-621.

Levy, M., and Y. Ritov. "Portfolio Optimization with Many Assets: The Importance of Short-Selling." Working paper, 2001.

Markowitz, H. M. "Portfolio Selection." *Journal of Finance*, vol. 7 (1952), pp. 77-91.

———. *Portfolio Selection: Efficient Diversification of Investments.* New Haven: Yale University Press, 1959.

Michaud, R. O. "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?" *Financial Analysts Journal*, vol. 45 (1989), pp. 31-42.

Pafka, S., M. Potters, and I. Kondor. "Exponential Weighting and Random-Matrix-Theory-Based Filtering of Financial Covariance Matrices for Portfolio Optimization." 2004. Available at http://arxiv.org/abs/cond-mat/0402573.

Sharpe, W. "A Simplified Model for Portfolio Analysis." *Management Science*, vol. 9, no. 1 (1963), pp. 277-293.

Stein, C. "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution." *Proceedings of the 3rd Berkeley Symposium on Probability and Statistics.* Berkeley: University of California Press, 1955, pp. 197-206.

Wolf, M. "Resampling vs. Shrinkage for Benchmarked Managers." *Wilmott Magazine*, vol. 1 (2007), pp. 76-82.