PORTFOLIO OPTIMIZATION USING A BLOCK STRUCTURE FOR THE COVARIANCE MATRIX

by

D. Disatnik

Portfolio Optimization Using a Block Structure for the Covariance Matrix

David Disatnik
The Leon Recanati Graduate School of Business Administration
Tel Aviv University
Tel Aviv, 69978, Israel
daviddis@post.tau.ac.il
+972-3-6408933

April 2010

Abstract

Implementing in practice the classical mean-variance theory for portfolio selection often results in obtaining portfolios with large short sale positions. Also, recent papers show that, due to estimation errors, existing and rather advanced mean-variance theory-based portfolio strategies do not consistently outperform the naïve $1/N$ portfolio that invests equally across N risky assets. In this paper, I introduce a portfolio strategy that generates a portfolio, with no short sale positions, that can outperform the $1/N$ portfolio. The strategy is investing in a global minimum variance portfolio (GMVP) that is constructed using an easy to calculate block structure for the covariance matrix of asset returns. Using this new block structure, the weights of the stocks in the GMVP can be found analytically, and as long as simple and directly computable conditions are met, these weights are positive.

Portfolio Optimization Using a Block Structure for the Covariance Matrix

I. Background

According to the seminal work of Markowitz (1952, 1959) an investor who cares only about the mean and variance of portfolio returns should hold a portfolio on the efficient frontier. Implementing in practice the mean-variance theory of Markowitz requires estimating the means and covariances of asset returns, and often results in obtaining portfolios with large short sale positions. This is true both when the means and covariances are estimated by the traditional sample mean vector and the traditional sample covariance matrix respectively, as well as by more advanced estimation techniques.[1]

Obtaining portfolios with short sale positions might be considered a drawback, since often short selling is restricted by regulators, in many cases investment policies of mutual funds prohibit taking short positions, and many individual investors find short selling onerous or impossible.[2] To the extent that short sales are indeed considered an undesirable feature of portfolio optimization, there is some interest in finding ways to produce efficient portfolios with long-only positions (henceforth—long-only portfolios).

Specifically, there is some interest in obtaining a long-only global minimum variance portfolio (henceforth—GMVP), which is, in the mean-variance framework, the portfolio on the efficient frontier with the smallest return variance.[3] The interest in the GMVP stems from the fact that several empirical studies show that in practice out-of- sample (ex-post) the GMVP performs at least as well as other frontier

---

[1] See for example: Ledoit and Wolf (2003), Jagannathan and Ma (2003), DeMiguel et al. (2009), and Disatnik and Benninga (2007).

[2] See for example regulation SHO of the U.S Securities and Exchange Commission (SEC), http://www.sec.gov/spotlight/shortsales.htm. Almazan et al. (2004) report that over the 1994-2000 period 69% of their sample U.S. domestic equity mutual funds were not allowed to short.

[3] Note that in practice the GMVP often includes less extreme short sale positions than other efficient portfolios (see for example Jagannathan and Ma [2003]). Yet, these short sale positions are still quite significant (see for example Disatnik and Benninga [2007]).

portfolios, even when the performance is evaluated based on measures related to both the return mean and variance (as opposed to the variance alone), such as the ex-post Sharpe ratio.[4] The common explanation for the relatively good ex-post performance of the GMVP is that the derivation of the GMVP requires estimating only the covariance matrix of asset returns, whereas for other efficient portfolios we have to estimate the means of asset returns as well, and that significantly adds to the estimation error.[5]

One way to obtain a long-only GMVP (as well as other long-only frontier portfolios) is by imposing on the optimization problem short sale constraints that prevent the portfolio weights from being negative (constrained optimization). Yet, at least from a theoretical point of view, this procedure is problematic, as it generates a portfolio whose weights can only be found numerically and not analytically.[6] Another issue with imposing the short sale constraints, as noted by Black and Litterman (1992), is that they generate "corner" solutions with zero weights in many assets, implying that the short sale constraints actually harm diversification.

In this paper, I introduce a new structure for the covariance matrix of asset returns—the block structure—which under simple and directly computable conditions generates (in an unconstrained optimization) a long-only GMVP whose weights can be found analytically. These conditions can maintain diversification, as they do not necessarily impose "corner" solutions with many zero weights.

To construct a block covariance matrix, one divides the portfolio's stocks into several groups (blocks). Within each block, the covariance between stocks is identical for all pairs of stocks in the block. The covariance between stocks from different blocks is also identical for all pairs. Thus, in the block structure, the number of covariances associated to each stock is reduced to two; the covariance with the

_____

[4] See for example: Jorion (1985, 1986, 1991), Jagannathan and Ma (2003), and DeMiguel et al. (2009).

[5] Jagannathan and Ma (2003), for instance, note that "the estimation error in the sample mean is so large that nothing much is lost in ignoring the mean altogether when no further information about the population mean is available."

[6] To obtain a solution for the constrained optimization problem, an iterative procedure, based on the Kuhn-Tucker conditions, is commonly used.

other stocks in the same block (the within-block covariance) and the covariance with the stocks from the other blocks (the between-block covariance).

I show that the weights of a GMVP constructed using the block matrix can be written as a function of the stock variances and covariances (namely, going further than the general solution for the unconstrained problem $\mathbf{w}_{GMVP} = \mathbf{\Sigma^{-1}1}/\mathbf{1^T\Sigma^{-1}1}$, where $\mathbf{\Sigma^{-1}}$ denotes the inverse matrix of the covariance matrix and $\mathbf{1}$ [$\mathbf{1^T}$] denotes a vector [a transpose vector] of ones). In essence, adding the conditions that: 1. the variance of each stock is greater than both its within-block covariance and its between-block covariance, and 2. the within-block covariances are not smaller than the between-block covariance is sufficient to ensure obtaining a long-only GMVP.

I find the block structure appealing from an implemental perspective. It requires the estimation only of the variances and a relatively small number of covariances, thus reducing the severe sampling error caused by having to estimate the whole covariance matrix. The condition that the variance of each stock is greater than its two associated covariances implies that we deal with relatively small covariances. Namely, like the shrinkage estimators, advocated by Ledoit and Wolf (2003, 2004), and the portfolios of estimators, advocated by Jagannathan and Ma (2000), Bengtsson and Holst (2002), and Disatnik and Benninga (2007), the block matrix has the appealing property of off-diagonal elements which are shrunk compared to the typically large off-diagonal elements of the traditional sample matrix. In addition, the sufficient conditions that ensure obtaining a long-only GMVP allow for nonnegative covariances—a robust characteristic of the U.S. stock market.[7]

I demonstrate empirically that even a rather simple example of the block matrix that I use for constructing the GMVP performs well. In a New York Stock Exchange (NYSE) dataset, the GMVP constructed using the block matrix outperforms the 1/N portfolio that invests equally across N risky assets, and which was recently highlighted by the comprehensive study of DeMiguel et al. (2009). They

---

[7] See for example: Chan et al. (1999) and the 2002 yearbook of Ibboston Associates.

show that, due to estimation errors, existing and rather advanced mean-variance theory-based portfolio strategies do not consistently outperform the naïve 1/N portfolio. In addition, Duchin and Levy (2009) show that that the 1/N portfolio outperforms relatively small mean-variance theory-based portfolios. My finding, on the other hand, may suggest that the mean-variance theory could be useful in practice, and is in line with Tu and Zhou (2009) who reach the same conclusion. The GMVP constructed using the block matrix also outperforms the value-weighted market portfolio from the CAPM world.

This study also continues previous studies that deal with analytical conditions related to long-only efficient portfolios. Rudd (1977), which corrects Roll (1977), goes in the opposite direction to mine and shows that if the GMVP is long-only and the inverse of the covariance matrix has positive diagonals and non-positive off-diagonals, then the variance of every individual stock is larger than each of its associated covariances. Roll and Ross (1977) bring a theorem that can be used to obtain conditions on a structure of a covariance matrix that generates a long-only GMVP. However, the conditions of this theorem are not computed as easily as mine. Kandel (1984) shows that for any set of N-1 assets, an Nth asset can be analytically constructed such that a long-only efficient portfolio will be obtained; however, as Levy and Ritov (2001) show, in large markets this N^{th} asset might be very unrealistic.

Green (1986) (which is slightly modified by Nielsen [1987]), by employing duality theory, presents necessary and sufficient conditions for efficient portfolios other than the GMVP to be long-only, as well as another set of necessary and sufficient conditions for obtaining a long-only GMVP. These conditions involve the feasibility of portfolios that have non-negative correlation with all assets and positive correlation with at least one. Also Best and Grauer (1992) derive general necessary and sufficient conditions for obtaining long-only efficient portfolios. They show that either there is no long-only efficient portfolio or there is a single segment of the efficient frontier for which all the portfolios are long-only. Best and Grauer's (1992) conditions are based on a scalar parameter that can be interpreted as an investor's risk tolerance parameter. A typical criticism regarding the conditions of Green (1986) and Best and Grauer (1992) might be that they are not always constructive.

The remainder of this paper proceeds as follows: First, I present the block structure. Then, I discuss some implemental aspects of the block structure and present the empirical illustration of its performance. I conclude the paper with a brief summary.

II. The block structure

I start with the special case of a two-block matrix. I assume a universe with n stocks. A covariance matrix Σ is said to be *two-block* if it has the following form:

$$
\Sigma =
\left(
\begin{array}{cccc|cccc}
s_1^2 & \eta_1 & \cdots & \eta_1 & \eta & \cdots & & \eta \\
\eta_1 & s_2^2 & \ddots & \vdots & \vdots & \ddots & & \vdots \\
\vdots & \ddots & \ddots & & & & & \\
 & & & \eta_1 & & & & \\
\eta_1 & \cdots & \eta_1 & s_j^2 & \eta & \cdots & & \eta \\
\hline
\eta & \cdots & & \eta & s_{j+1}^2 & \eta_2 & \cdots & \eta_2 \\
\vdots & \ddots & & \vdots & \eta_2 & s_{j+2}^2 & \ddots & \vdots \\
 & & & & \vdots & \ddots & \ddots & \\
 & & & & & & & \eta_2 \\
\eta & \cdots & & \eta & \eta_2 & \cdots & \eta_2 & s_n^2 \\
\end{array}
\right)
$$

Here j and $n\text{-}j$ are the sizes of the two blocks (j and $n\text{-}j$ are not necessarily equal), s_i^2 are the variances, η_1 and η_2 are the within-block covariances, and η is the between-block covariance.

Proposition 1 below characterizes sufficient conditions on η_1, η_2 and η under which the two-block matrix produces (in an unconstrained optimization) a long-only GMVP. Note that the proposition's conditions guarantee not only obtaining a long-only GMVP but also that the two-block matrix is an invertible covariance matrix. Without further restricting η_1, η_2 and η, the conditions derived are a bit messy. Yet, restricting η_1, η_2 and η to be nonnegative generates simple and directly computable conditions: 1. the variance of each stock is greater than both its within-block covariance and its between-block covariance, and 2. the within-block covariances are not smaller than the between-block covariance.

Proposition 1: Suppose that Σ is a two-block matrix. Then Σ produces a long-only GMVP if the following conditions on η_1, η_2 and η hold:

$$-\left|\eta_1^*\right| < \eta_1 < \min\left(s_i^2\right) \quad , \quad i = 1,...,j$$

$$-\left|\eta_2^*\right| < \eta_2 < \min\left(s_i^2\right) \quad , \quad i = j+1,...,n$$

$$-\left|\eta_{12}^*\right| < \eta \leq \min\left(\eta_1,\eta_2\right)$$

where $\left|\eta_1^*\right|$ and $\left|\eta_2^*\right|$ are respectively the unique solutions of the following equations:

$$\left|\eta_1^*\right| = 1 \Big/ \sum_{i=1}^{j} \frac{1}{s_i^2 + \left|\eta_1^*\right|} \quad \text{and} \quad \left|\eta_2^*\right| = 1 \Big/ \sum_{i=j+1}^{n} \frac{1}{s_i^2 + \left|\eta_2^*\right|}$$

and:

$$\left|\eta_{12}^*\right| = +\sqrt{\left(\eta_1 + \frac{1}{\displaystyle\sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1}}\right)\left(\eta_2 + \frac{1}{\displaystyle\sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}}\right)}$$

Proposition 1 is proved in the appendix to the paper.

When η_1, η_2 and η are restricted to be nonnegative, we get the following simple and directly computable sufficient conditions for which the GMVP is long-only:

Corollary: Suppose that Σ is a two-block matrix and the covariances η_1, η_2 and η are nonnegative. Then Σ produces a long-only GMVP if the following conditions on η_1, η_2 and η hold:

$$0 \leq \eta_1 < \min\left(s_i^2\right) \quad , \quad i = 1,...,j$$

$$0 \leq \eta_2 < \min\left(s_i^2\right) \quad , \quad i = j+1,...,n$$

$$0 \leq \eta \leq \min\left(\eta_1,\eta_2\right)$$

As shown in the proof of Proposition 1, the expressions for the weight in the GMVP of a stock from the first and the second block are respectively:

$$w_i = \frac{1}{s_i^2 - \eta_1} \cdot \frac{1 + (\eta_2 - \eta)A_2}{A_1 + A_2 + (\eta_1 + \eta_2 - 2\eta)A_1 A_2} \quad , \quad i = 1, ..., j$$

$$\quad , \quad (1)$$

$$w_i = \frac{1}{s_i^2 - \eta_2} \cdot \frac{1 + (\eta_1 - \eta)A_1}{A_1 + A_2 + (\eta_1 + \eta_2 - 2\eta)A_1 A_2} \quad , \quad i = j+1, ..., n$$

where: $A_1 = \sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1}$, $A_2 = \sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}$.

Assuming that the conditions for having a long-only GMVP hold, we can see that in each block the weights in the GMVP are biased towards the stocks with the relatively small variances. However, overall, the stocks with the smallest variances are not necessarily also those with the largest weights in the GMVP. That is because two more effects should be taken into consideration. The first is the differences between the stock variances and the within-block covariances. The second is the differences between the within-block covariances and the between-block covariance.[8]

Two special cases are: When all the covariances equal zero, we get the diagonal covariance matrix, for which the GMVP weights are biased towards the stocks with the smaller variances. When all the covariances equal a constant other than zero, the differences between the stock variances and the constant covariance determine the GMVP weights, implying that overall the GMVP weights are biased towards the stocks with the relatively small variances. We can also see that the conditions of Proposition 1 are flexible enough to generate positive weights without harming diversification.

---

[8] Assume that we have four stocks, two in each block. In the first block, $s_1^2 = 0.35$, $s_2^2 = 0.37$, $\eta_1 = 0.27$; in the second block, $s_3^2 = 0.4$, $s_4^2 = 0.5$, $\eta_2 = 0.01$; and the third covariance is $\eta = 0$. The GMVP weights here are: $w_1 = 0.23$, $w_2 = 0.19$, $w_3 = 0.32$, $w_4 = 0.26$ —a larger fraction of the GMVP is invested in the stocks with the larger variances.

I now turn to the general block matrix. I assume a universe with n stocks. A covariance matrix $\boldsymbol{\Sigma}$ is said to be a *block matrix* if it has the following form:

$$\boldsymbol{\Sigma} = \begin{pmatrix} s_1^2 & \eta_1 & \cdots & \eta_1 & & & & & & & & \\ \eta_1 & \ddots & \ddots & \vdots & & \eta & & & & & \eta & \\ \vdots & \ddots & \ddots & \eta_1 & & & & & & & & \\ \eta_1 & \cdots & \eta_1 & s_k^2 & & & & & & & & \\ & & & & s_{k+1}^2 & \eta_2 & \cdots & & \eta_2 & & & \\ & & & & \eta_2 & \ddots & \ddots & & \vdots & & & \\ & & & & \vdots & \ddots & s_l^2 & & & & & \\ & \eta & & & & & \ddots & & & & \eta & \\ & & & & & & & \eta_2 & & & & \\ & & & & \eta_2 & \cdots & & \eta_2 & s_p^2 & & & \\ & & & & & & & & & \ddots & & \\ & & & & & & & & & s_r^2 & \eta_M & \cdots & \eta_M \\ & & & & & & & & & \eta_M & \ddots & \\ & \eta & & & & \eta & & & & \vdots & \ddots & s_t^2 \\ & & & & & & & & & & & \ddots & \eta_M \\ & & & & & & & & \eta_M & & & \eta_M & s_n^2 \end{pmatrix}$$

Here we have M blocks of stocks (which are not required to include an equal number of stocks), s_j^2, $j=1,\ldots,n$, are the sample variances, η_i, $i=1,\ldots,M$, are the within-block covariances, and η is the between-block covariance.

Proposition 2 below characterizes sufficient conditions on η_i and η under which the block matrix produces (in an unconstrained optimization) a long-only GMVP. I only present here the simple and directly computable conditions that are obtained when all the covariances are restricted to be nonnegative. Like in the special case of the two-block matrix, also for the general block matrix the sufficient conditions are: 1. the variance of each stock is greater than both its within-block covariance and its between-block covariance, and 2. the within-block covariances are not smaller than the between-block covariance.

Proposition 2: Suppose that Σ is a block matrix and the covariances η_i, $i=1,...M$ and η are nonnegative. Then Σ produces a long-only GMVP if the following conditions on η_i and η hold:

$$0 \le \eta_i < \text{minvar}_i \quad , \quad i=1,....,M$$
$$0 \le \eta \le \min\left(\eta_i\right) \quad , \quad i=1,....,M$$

where minvar_i denotes the minimal variance in block i. Proposition 2 is proved in the appendix to the paper.

As shown in the proof of Proposition 2, the expression for the weight in the GMVP of stock j from block i ($i=1,…,M$; $j=1,...,$ Z_i, where Z_i denotes the number of stocks in block i) is:

$$w_j = \frac{1}{s_j^2 - \eta_i} \cdot \frac{1 + \sum_{k=1;k\neq i}^{M} B_k + \sum_{k,l=1;k<l;k,l\neq i}^{M} B_k B_l + \sum_{k,l,r=1;k<l<r;k,l,r\neq i}^{M} B_k B_l B_r + ... + \underset{k<l<r<...t...<M;k,l,r...,t...,M\neq i}{B_k B_l B_r...B_t...B_M}}{\sum_{i=1}^{M} A_i C_i} \quad , (2)$$

where: $A_i = \sum_{j=1}^{Z_i} \frac{1}{s_j^2 - \eta_i}$, $B_i = \left(\eta_i - \eta\right) A_i$, $i=1,...,M$,

and $\quad C_i = 1 + \sum_{k=1;k\neq i}^{M} B_k + \sum_{k,l=1;k<l;k,l\neq i}^{M} B_k B_l + \sum_{k,l,r=1;k<l<r;k,l,r\neq i}^{M} B_k B_l B_r + ... + \underset{k<l<r<...t...<M;k,l,r...,t...,M\neq i}{B_k B_l B_r...B_t...B_M}$.[9]

Like in the special case of the two-block matrix, also here, when the conditions for having a long-only GMVP hold, in each block the weights in the GMVP are biased towards the stocks with the relatively small variances. As before, also now, overall, the stocks with the smallest variances are not necessarily also those with the largest weights in the GMVP, since the differences between the stock variances and the within-block covariances, and the differences between the within-block covariances and the between-block covariance should also be taken

---

[9] I show in the appendix to the paper that expression (1) is indeed a specific case, $M=2$, of the general expression (2).

into consideration. Also with the general block matrix, we can have positive weights and maintain diversification.

III. Implemental aspects—the block matrix

In the previous section, I presented the block matrix. In this section, I explain why I find this structure appealing from an implemental perspective, and present an empirical illustration of its performance using historical data of NYSE stock returns.

First, the block structure requires the estimation only of the variances and a relatively small number of covariances, thus reducing the severe sampling error caused by having to estimate the whole covariance matrix (like in the traditional sample matrix case). As discussed in Disatnik and Benninga (2007), reducing the sampling error usually gives rise to specification error, since it is often done by imposing some form of structure on the estimation model that is being used. Therefore, in essence, the real goal from an estimation point of view is not just reducing the sampling error, but to find the covariance matrix structure that can create the optimal tradeoff between the sampling error and the specification error. The block structure might fit for this purpose, since it leaves enough degrees of freedom in respect to the number of the blocks used and the number of the stocks included in each block.

Second, the sufficient conditions that ensure obtaining a long-only GMVP involve relatively small covariances (compared to the variances on the diagonal). Namely, like the shrinkage and the portfolios of estimators mentioned before, the block matrix has the appealing property of off-diagonal elements which are shrunk compared to the typically large off-diagonal elements of the traditional sample matrix. This property is appealing, because the large off-diagonal elements are those responsible for the extreme long and short positions that are obtained so often when the mean-variance theory is implemented in practice. Michaud (1989) also states that inverting the sample matrix with its large off-diagonal elements (as required by the mean-variance theory) even amplifies the sampling error.

Third, the sufficient conditions that ensure obtaining a long-only GMVP allow for nonnegative covariances, which is appealing since generally, as mentioned before, the covariances between stocks in the U.S. stock market are positive.

To illustrate, I evaluate empirically the performance of a rather unsophisticated example of the block matrix. I divide the stocks into five blocks based on their market capitalization. I use the sample variances to estimate the variances on the diagonal, set the within-block covariance in each block equal to 95 percent of the smallest estimated variance in that block, and set the between-block covariance to zero.

I evaluate the performance of the specific five-block matrix by comparing the ex-post Sharpe ratio of a GMVP constructed using the specific block matrix, with the ex-post Sharpe ratio of a GMVP constructed using the diagonal sample matrix (which can be viewed as another specific block matrix; covariances set to zero and sample variances on the diagonal), and with the ex-post Sharpe ratio of the "naïve" $1/N$ portfolio, in which a fraction $1/N$ of wealth is invested in each of the N stocks available.

I use the $1/N$ portfolio following the recent comprehensive study of DeMiguel et al. (2009), in which they evaluate the performance of fourteen models of portfolio selection across seven different datasets. Their main finding is that none of the fourteen evaluated models, including rather advanced models, generates consistently higher ex-post Sharpe Ratios than the naïve $1/N$ portfolio.[10] Duchin and Levy (2009) show that that the $1/N$ portfolio outperforms relatively small mean-variance theory-based portfolios. As a result, obtaining in my illustration that the example of the block matrix performs at least as well as the $1/N$ portfolio will enable to conclude that (at least for my NYSE dataset) the block construct can perform well. The optimal strategy in a CAPM world is the value-weighted market portfolio. So, in the demonstration, I also use a benchmark value-weighted portfolio and report its ex-post Sharpe ratio.

---

[10] The advanced techniques evaluated in DeMiguel et al. (2009) include Bayesian models for estimation, the MacKinlay and Pastor's (2000) missing factor model, the Kan and Zhou's (2007) three-fund model, and the multi-prior model of Garlappi et al. (2006).

The block matrix is being used to construct a GMVP. Therefore, I also evaluate its performance using a comparison of the ex-post standard deviation of the GMVP with the ex-post standard deviations of the other three portfolios in the demonstration.

Following Chan et al. (1999), Bengtsson and Holst (2002), Jagannathan and Ma (2003), Ledoit and Wolf (2003), and DeMiguel et al. (2009), I use the "rolling-sample" approach for calculating the ex-post Sharpe ratios and standard deviations in the demonstration.

In a nutshell, to implement the "rolling-sample" approach, given a T-month long dataset of monthly stock returns, I choose an estimation window (in-sample period) of length L months. In month t=L+1, I use the returns in the previous L months to estimate the parameters needed to each of my two covariance matrix estimators. Each of the covariance matrix estimators are then used to determine the weights in the two corresponding GMVP. These weights are then used to compute the monthly returns on each of the two GMVP in the out-of-sample period from t=L+1 till t=L+k. In month t=L+k+1, I start the whole process all over again; namely, using the returns in the previous L months to estimate the two covariance matrix estimators, determining the weights in the two corresponding GMVP, and computing the monthly returns on the two GMVP in the out-of-sample period of k months. The process is continued by adding the returns of the next k months in the dataset and dropping the earliest k returns, until the end of the dataset is reached. The outcome of this approach is a series of *T-L* monthly out-of-sample (ex-post) returns generated by each of the two covariance matrix estimators.[11]

In the same manner, but of course without a need to estimate the covariance matrix, I derive the series of the ex-post returns on the 1/*N* portfolio and the value-weighted portfolio. For the value-weighted portfolio, in each rebalancing point, I calculate its weights based on the market capitalization of the stocks included in the dataset.

---

[11] The description of the "rolling-sample" approach relies on DeMiguel et al. (2009).

I subtract from each of the four ex-post return series the corresponding *T-L* one-month T-bill returns, which are extracted from Ken French's website.[12] By that I obtain four series of ex-post excess returns, which enable to compute the corresponding ex-post standard deviations and Sharpe ratios. For series *i*, the ex-post Sharpe ratio, SR_i, is defined as $SR_i = Z_i/\sigma_i$, where Z_i and σ_i denote respectively the ex-post expected excess return and the ex-post standard deviation of series *i*.

I conduct the empirical evaluation six times, each time changing the length of the in-sample period or the length of the out-of-sample period. I use in-sample periods of 120 months (also used in Ledoit and Wolf [2003]) and 60 months (also used in Chan et al. [1999] and Jagannathan and Ma [2003]).[13] I use out-of-sample periods of 12 months (also used in Chan et al. [1999], Jagannathan and Ma [2003], and Ledoit and Wolf [2003]), 24 months, and 36 months. I chose these three out-of-sample periods, since I believe they correspond to realistic investment horizons (see also Chan et al. [1999]). I use monthly returns on stocks traded on the NYSE. The stock returns are extracted from the Center for Research in Security Prices (CRSP) database. The period of the study (*T*) is from 1/1964 to 12/2003.

As an aside, each time I construct the portfolios, I do it only out of NYSE stocks whose returns cover the entire in-sample and out-of-sample periods used. For example, in the case of in-sample period of 120 months and out-of sample period of 12 months, for constructing the GMVP of 1/74, I only use NYSE stocks with monthly return data for all the 132 months from 1/64 till 12/74. For constructing the GMVP of 1/75, I only use NYSE stocks with monthly return data for all the 132 months from 1/65 till 12/75 and so on (see also Bengtsson and Holst [2002]).[14]

---

[12] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[13] Jobson and Korkie (1981) mention rules of thumb regarding the length of the in-sample period of 4 to 7 years and 8 to 10 years.

[14] I am aware of the fact that this widely-followed procedure introduces survivorship bias into the estimation procedure. However, since the survivorship bias is common to all the compared estimators, I do not consider this a significant problem.

The table below presents the ex-post Sharpe ratios and standard deviations obtained in the demonstration. The Sharpe ratios and the standard deviations are annualized through multiplication by $\sqrt{12}$.

Ex-Post Sharpe Ratios						
	In sample 120 months			In sample 60 months		
	Out of sample period					
	12 months	24 months	36 months	12 months	24 months	36 months
My specific five-block matrix	0.67	0.66	0.65	0.68	0.65	0.63
Diagonal sample matrix	0.65	0.65	0.64	0.64	0.64	0.63
1/N portfolio	0.60	0.60	0.60	0.59	0.57	0.60
Value-weighted portfolio	0.44	0.45	0.46	0.43	0.44	0.45
Ex-Post Standard Deviations						
	In sample 120 months			In sample 60 months		
	Out of sample period					
	12 months	24 months	36 months	12 months	24 months	36 months
My specific five-block matrix	11.49%	11.51%	11.38%	11.36%	11.40%	11.26%
Diagonal sample matrix	13.17%	13.17%	13.21%	12.95%	12.97%	12.88%
1/N portfolio	16.24%	16.18%	16.11%	16.62%	16.38%	16.31%
Value-weighted portfolio	15.21%	15.32%	15.35%	15.38%	15.49%	15.53%

This table reports the annualized ex-post Sharpe ratios and standard deviations generated by the two tested covariance matrix estimators, the 1/N portfolio, and the value-weighted portfolio in the six runs of the demonstration.

In the six runs conducted, the GMVP constructed using the five-block matrix generates the highest ex-post Sharpe ratios and the lowest ex-post standard deviations.[15] Together with the findings of DeMiguel et al. (2009) and Duchin and Levy (2009) that rather advanced portfolio selection models do not perform consistently better than the 1/N portfolio, my demonstration may suggest that the block method could have an impact on practical portfolio choice. Future research should address the practical implementation of the block construct more extensively. Specifically, the following issues should be examined: Which criteria to use for dividing the stocks into the blocks; how many blocks to use; how to estimate the within-block covariances; and how to estimate the between-block covariance.

---

[15] Note that also the GMVP constructed using the diagonal sample matrix always generates higher ex-post Sharpe ratios and lower standard deviations than the 1/N portfolio and the benchmark value-weighted portfolio.

IV. Summary

To the extent that short sales are considered an undesirable feature of portfolio optimization, there is some interest in discussing long-only efficient portfolios. Specifically, there is some interest in finding estimators of the covariance matrix that generate a long-only GMVP, as several empirical studies show that ex-post the GMVP performs at least as well as other frontier portfolios. Imposing short sale constraints on the optimization problem, no matter which covariance matrix estimator is used, enables to obtain a long-only GMVP. Yet, at least from a theoretical point of view, this way is problematic, as it produces a GMVP whose weights can only be found numerically and not analytically. Imposing the short sale constraints also generates "corner" solutions with zero weights in many assets, implying that the short sale constraints actually harm diversification.

In this paper, I introduce a new structure for the covariance matrix—the block structure—which under simple and directly computable conditions generates (in an unconstrained optimization) a long-only GMVP whose weights can be found analytically. With these conditions, having positive weights can come not at the expense of diversification.

To construct a block covariance matrix, one divides the portfolio's stocks into several groups (blocks). Within each block, the covariance between stocks is identical for all pairs of stocks in the block. The covariance between stocks from different blocks is also identical for all pairs. I show that the weights of a GMVP constructed using the block matrix can be found analytically. These weights are positive as long as the variance of each stock is greater than both its within-block covariance and its between-block covariance, and as long as the within-block covariances are not smaller than the between-block covariance.

I find the block structure appealing from an implemental perspective, as it consists of a relatively small number of relatively small covariances, which can be nonnegative. I show empirically that a rather simple example of the block matrix that I use for constructing a GMVP performs well. The GMVP outperforms the 1/*N* portfolio which was highlighted by recent papers that find that existing and rather advanced mean-variance theory-based portfolio strategies do not consistently outperform the 1/*N*

portfolio. Thus, my finding may suggest that the block structure for the covariance matrix of asset returns could have an impact on practical portfolio choice. Future research should address more extensively the usage of the block matrix in practice.

References

Almazan, A., Brown, K.C., Carlson, M., and D.A. Chapman. (2004). Why Constrain your Mutual Fund Manager? *Journal of Financial Economics,* 73: 289-321.

Bengtsson, C. and J. Holst. (2002). On Portfolio Selection: Improved Covariance Matrix Estimation for Swedish Asset Returns. *Working paper,* Lund University.

Best, M.J and R. R. Grauer. (1992). Positively Weighted Minimum-Variance Portfolios and the Structure of Asset Expected Returns. *Journal of Financial and Quantitative Analysis*, 27(4): 513-537.

Black, F. and R. Litterman. (1992). Global Portfolio Optimization. *Financial Analysis Journal*, 48(5): 28-43.

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

DeMiguel, V., L. Garlappi, and R. Uppal. (2009). Optimal versus Naïve Diversification: How Inefficient is the 1/*N* Portfolio Strategy? *Review of Financial Studies*, 22(5): 1915-1953.

Disatnik, D.J. and S. Benninga. (2007). Shrinking the Covariance Matrix. *Journal of Portfolio Management,* 33(4): 55-63.

Duchin, R. and H. Levy [2009]. Markowitz versus the Talmudic Portfolio Diversification Strategies. *Journal of Portfolio Management*, 35(2): 71-74.

Garlappi, L., R. Uppal and T. Wang. (2006). Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. Forthcoming in *Review of Financial Studies.*

Green, R.C. (1986). Positively Weighted Portfolios on the Minimum-Variance Frontier. *Journal of Finance*, 41(5): 1051-1068.

Ibboston Associates, Chicago. (2002). *Stocks, Bonds, Bills, and Inflation 2002 Yearbook.*

Jagannathan, R. and T. Ma. (2000). Three Methods for Improving the Precision in Covariance Matrix Estimation. *Unpublished working paper*.

Jagannathan, R. and T. Ma. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance,* 54(4): 1651-1683.

Jobson, J. D. and B. Korkie. (1981). Putting Markovitz Theory to Work. *Journal of Portfolio Management,* 7: 70-74.

Jorion, P. (1985). International Portfolio Diversification with Estimation Risk. *Journal of Business,* 58: 259-278.

Jorion, P. (1986). Bayes-Stein Estimation for Portfolio Analysis. *Journal of Financial and Quantitative Analysis,* 21: 279-292.

Jorion, P. (1991). Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection. *Journal of Banking and Finance,* 15: 717-727.

Kan, R. and G. Zhou. (2007). Optimal Portfolio Choice with Parameter Uncertainty. *Journal of Financial and Quantitative Analysis*, 42: 621-656.

Kandel, S. (1984). On the Exclusion of Assets from Tests of the Mean Variance Efficiency of the Market Portfolio. *Journal of Finance*, 39: 63-75.

Ledoit, O. and M. Wolf. (2003). Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection. *Journal of Empirical Finance,* 10: 603-621.

Ledoit, O. and M. Wolf. (2004). Honey, I Shrunk the Sample Covariance Matrix. *Journal of Portfolio Management,* 30(4): 110-119.

Levy, M. and Y. Ritov. (2001). Portfolio Optimization with Many Assets: The Importance of Short-Selling. *Working Paper,* the Hebrew University of Jerusalem.

MacKinlay, A.C. and L. Pastor. (2000). Asset Pricing Models: Implications for Expected returns and Portfolio Selection. *Review of Financial Studies*, 13: 883-916.

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

Nielsen, L.T. (1987). Positive Weighted Frontier Portfolios: A Note. *Journal of Finance*, 42(2): 471.

Roll, R. (1977). A Critique of the Asset Pricing Theory's Tests; Part I: On Past and Potential Testability of the Theory. *Journal of Financial Economics*, 4: 129-176.

Roll, R. and S.A. Ross. (1977). Comments on Qualitative Results for Investment Proportions. *Journal of Financial Economics*, 5:265-268.

Rudd, A. (1977). A Note on Qualitative Results for Investment Proportions. *Journal of Financial Economics*, 5:259-263.

Tu, J. and G. Zhou. (2009). Markowitz Meets Talmud: A Combination of Sophisticated and Naive Diversification Strategies. *Working paper,* Singapore Management University and Washington University at St. Louis.

Appendix: Proofs

The proof of Proposition 1

The proof consists of two phases:

1. Finding sufficient conditions on η_1, η_2 and η for which a long-only GMVP is obtained.

2. Modifying the conditions to guarantee that the two-block matrix is an invertible covariance matrix. Namely, showing that the eigenvalues of the matrix are strictly positive, since every positive definite symmetric matrix is an invertible covariance matrix.

Phase 1—long-only GMVP

$$\text{Denote the first column of } \mathbf{\Sigma^{-1}} \text{ by:} \quad \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix}$$

where $\mathbf{\Sigma^{-1}}$ denotes the inverse matrix of $\mathbf{\Sigma}$.

Then, since $\mathbf{\Sigma}\mathbf{\Sigma^{-1}} = \mathbf{I}$:

$$\mathbf{\Sigma} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \\ \\ \\ 0 \end{pmatrix}$$

Given the structure of Σ, writing this explicitly yields:

$$s_1^2 x_1 + \eta_1 x_2 + \ldots + \eta_1 x_j + \eta x_{j+1} + \eta x_{j+2} + \ldots + \eta x_n = 1$$

$$\eta_1 x_1 + s_2^2 x_2 + \ldots + \eta_1 x_j + \eta x_{j+1} + \eta x_{j+2} + \ldots + \eta x_n = 0$$

$$\vdots$$

(A1)
$$\eta_1 x_1 + \eta_1 x_2 + \ldots + s_j^2 x_j + \eta x_{j+1} + \eta x_{j+2} + \ldots + \eta x_n = 0$$

$$\eta x_1 + \eta x_2 + \ldots + \eta x_j + s_{j+1}^2 x_{j+1} + \eta_2 x_{j+2} + \ldots + \eta_2 x_n = 0$$

$$\eta x_1 + \eta x_2 + \ldots + \eta x_j + \eta_2 x_{j+1} + s_{j+2}^2 x_{j+2} + \ldots + \eta_2 x_n = 0$$

$$\vdots$$

$$\eta x_1 + \eta x_2 + \ldots + \eta x_j + \eta_2 x_{j+1} + \eta_2 x_{j+2} + \ldots + s_n^2 x_n = 0$$

And therefore:

$$\left(s_1^2 - \eta_1\right)x_1 + \eta_1 \sum_{i=1}^{j} x_i + \eta \sum_{i=j+1}^{n} x_i = 1$$

$$\left(s_2^2 - \eta_1\right)x_2 + \eta_1 \sum_{i=1}^{j} x_i + \eta \sum_{i=j+1}^{n} x_i = 0$$

$$\vdots$$

$$\left(s_j^2 - \eta_1\right)x_j + \eta_1 \sum_{i=1}^{j} x_i + \eta \sum_{i=j+1}^{n} x_i = 0$$

$$\eta \sum_{i=1}^{j} x_i + \left(s_{j+1}^2 - \eta_2\right)x_{j+1} + \eta_2 \sum_{i=j+1}^{n} x_i = 0$$

$$\eta \sum_{i=1}^{j} x_i + \left(s_{j+2}^2 - \eta_2\right)x_{j+2} + \eta_2 \sum_{i=j+1}^{n} x_i = 0$$

$$\vdots$$

$$\eta \sum_{i=1}^{j} x_i + \left(s_n^2 - \eta_2\right)x_n + \eta_2 \sum_{i=j+1}^{n} x_i = 0$$

Dividing the first j equations by $s_i^2 - \eta_1$ (assuming $\eta_1 \neq s_i^2$, $i = 1,\ldots, j$) and dividing the last n-j

equations by $s_i^2 - \eta_2$ (assuming $\eta_2 \neq s_i^2$, $i = j+1,\ldots, n$) give:

$$x_1 + \frac{\eta_1}{s_1^2 - \eta_1} \sum_{i=1}^{j} x_i + \frac{\eta}{s_1^2 - \eta_1} \sum_{i=j+1}^{n} x_i = \frac{1}{s_1^2 - \eta_1}$$

$$\vdots$$

$$x_j + \frac{\eta_1}{s_j^2 - \eta_1} \sum_{i=1}^{j} x_i + \frac{\eta}{s_j^2 - \eta_1} \sum_{i=j+1}^{n} x_i = 0$$

(A2)

$$x_{j+1} + \frac{\eta}{s_{j+1}^2 - \eta_2} \sum_{i=1}^{j} x_i + \frac{\eta_2}{s_{j+1}^2 - \eta_2} \sum_{i=j+1}^{n} x_i = 0$$

$$\vdots$$

$$x_n + \frac{\eta}{s_n^2 - \eta_2} \sum_{i=1}^{j} x_i + \frac{\eta_2}{s_n^2 - \eta_2} \sum_{i=j+1}^{n} x_i = 0$$

Summing the first j equations in (A2) and rearranging terms give:

(A3) $$\left(1 + \eta_1 \sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1}\right) \sum_{i=1}^{j} x_i + \eta \sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1} \sum_{i=j+1}^{n} x_i = \frac{1}{s_1^2 - \eta_1}$$

Summing the last n-j equations in (A2) and rearranging terms give:

(A4) $$\sum_{i=j+1}^{n} x_i = -\frac{\eta \sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}}{1 + \eta_2 \sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}} \sum_{i=1}^{j} x_i$$

And now substituting (A4) into (A3) and rearranging terms give:

$$\sum_{i=1}^{j} x_i = \frac{1}{s_1^2 - \eta_1} \cdot \frac{1 + \eta_2 \sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}}{\left(1 + \eta_1 \sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1}\right)\left(1 + \eta_2 \sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}\right) - \eta^2 \sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1} \sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}}$$

Denote:

$$\Delta = \left(1 + \eta_1 \sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1}\right)\left(1 + \eta_2 \sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}\right) - \eta^2 \sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1} \sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}$$

And therefore:

$$\text{(A5)} \qquad \sum_{i=1}^{j} x_i = \frac{1}{s_1^2 - \eta_1} \cdot \frac{1 + \eta_2 \displaystyle\sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}}{\Delta}$$

Substituting (A5) into (A4) gives:

$$\text{(A6)} \qquad \sum_{i=j+1}^{n} x_i = \frac{1}{s_1^2 - \eta_1} \cdot \frac{-\eta \displaystyle\sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}}{\Delta}$$

Adding equations (A5) and (A6) gives the sum of (the elements in) the first column of $\mathbf{\Sigma^{-1}}$. Since $\mathbf{\Sigma^{-1}}$ is symmetric (because $\mathbf{\Sigma}$ is symmetric) this is also the sum of (the elements in) the first row of $\mathbf{\Sigma^{-1}}$:

$$\text{(A7)} \qquad \sum_{j=1}^{n} \Sigma^{-1}_{1j} = \frac{1}{s_1^2 - \eta_1} \cdot \frac{1 + (\eta_2 - \eta) \displaystyle\sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}}{\Delta}$$

Generalizing the last result and finding the sum of each one of the rows of $\mathbf{\Sigma^{-1}}$ is done as follows: I start with the first j rows of $\mathbf{\Sigma^{-1}}$. I repeat the above procedure [i.e., deriving equations (A1)-(A7)] j times. Each time the only difference is that the only row in (A1) that equals 1 moves one place ahead. For example, for the sum of the second row of $\mathbf{\Sigma^{-1}}$, the second row in (A1), and not the first row, equals 1. Thus, the general expression for the sum of one of the first j rows of $\mathbf{\Sigma^{-1}}$ is:

$$\text{(A8)} \qquad \sum_{j=1}^{n} \Sigma^{-1}_{ij} = \frac{1}{s_i^2 - \eta_1} \cdot \frac{1 + (\eta_2 - \eta) \displaystyle\sum_{i=j+1}^{n} \frac{1}{s_i^2 - \eta_2}}{\Delta}$$

Repeating a similar procedure $n-j$ times and applying "symmetric considerations" enable to find the general expression for the sum of one of the last $n-j$ rows of $\mathbf{\Sigma^{-1}}$:

$$\text{(A9)} \qquad \sum_{j=1}^{n} \Sigma^{-1}_{ij} = \frac{1}{s_i^2 - \eta_2} \cdot \frac{1 + (\eta_1 - \eta) \displaystyle\sum_{i=1}^{j} \frac{1}{s_i^2 - \eta_1}}{\Delta}$$

The vector of the GMVP weights is $\mathbf{w} = \dfrac{\mathbf{\Sigma^{-1}1}}{\mathbf{1^T\Sigma^{-1}1}}$. For the weight of stock i, in the numerator we

have the sum of row i in $\mathbf{\Sigma^{-1}}$ and in the denominator we have the sum of (all the elements in) $\mathbf{\Sigma^{-1}}$.

Hence, in order to obtain the denominator, I use (A8) and (A9):

$$\mathbf{1^T\Sigma^{-1}1} = \sum_{i=1}^{j}\left(\frac{1}{s_i^2 - \eta_1} \cdot \frac{1 + (\eta_2 - \eta)\sum_{i=j+1}^{n}\frac{1}{s_i^2 - \eta_2}}{\Delta} \right) + \sum_{i=j+1}^{n}\left(\frac{1}{s_i^2 - \eta_2} \cdot \frac{1 + (\eta_1 - \eta)\sum_{i=1}^{j}\frac{1}{s_i^2 - \eta_1}}{\Delta} \right)$$

And after rearranging a bit more we get:

(A10) $\quad \mathbf{1^T\Sigma^{-1}1} = \dfrac{\displaystyle\sum_{i=1}^{j}\frac{1}{s_i^2 - \eta_1} + \sum_{i=j+1}^{n}\frac{1}{s_i^2 - \eta_2} + (\eta_1 + \eta_2 - 2\eta)\sum_{i=1}^{j}\frac{1}{s_i^2 - \eta_1}\sum_{i=j+1}^{n}\frac{1}{s_i^2 - \eta_2}}{\Delta}$

Substituting (A8), (A9) and (A10) into $\mathbf{w} = \dfrac{\mathbf{\Sigma^{-1}1}}{\mathbf{1^T\Sigma^{-1}1}}$ gives the expressions for the weight in the

GMVP of a stock from the first and the second block respectively:

(A11)
$$w_i = \frac{1}{s_i^2 - \eta_1} \cdot \frac{1 + (\eta_2 - \eta)A_2}{A_1 + A_2 + (\eta_1 + \eta_2 - 2\eta)A_1A_2} \quad , \quad i = 1,...,j$$

$$w_i = \frac{1}{s_i^2 - \eta_2} \cdot \frac{1 + (\eta_1 - \eta)A_1}{A_1 + A_2 + (\eta_1 + \eta_2 - 2\eta)A_1A_2} \quad , \quad i = j+1,...,n$$

where: $A_1 = \displaystyle\sum_{i=1}^{j}\frac{1}{s_i^2 - \eta_1} \quad , \quad A_2 = \sum_{i=j+1}^{n}\frac{1}{s_i^2 - \eta_2}$

And now we can easily see that the vector $\mathbf{w}$ is strictly positive if the following set of conditions holds:

(A12)
$$\eta_1 < \min\left(s_i^2\right) \ , \ i = 1,...,j$$
$$\eta_2 < \min\left(s_i^2\right) \ , \ i = j+1,...,n$$
$$\eta \le \min\left(\eta_1, \eta_2\right)$$

Phase 2—strictly positive eigenvalues

$\boldsymbol{\Sigma}$ is a symmetric real matrix. Therefore, it has real eigenvalues and it can be diagonalized:

$$\mathbf{R^{-1}\Sigma R} = \boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ 0 & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

where $\mathbf{R}$ denotes the diagonalizing matrix, $\mathbf{R^{-1}}$ denotes the inverse matrix of $\mathbf{R}$, and $\boldsymbol{\Lambda}$ denotes the diagonal matrix, whose diagonal elements $\lambda_1, ..., \lambda_n$ are the eigenvalues of $\boldsymbol{\Sigma}$.

Hence, $\boldsymbol{\Sigma R} = \mathbf{R\Lambda}$ and the element ij of $\boldsymbol{\Sigma R}$ is given by:

$$\left(\boldsymbol{\Sigma R}\right)_{ij} = \sum_{k=1}^{n} \boldsymbol{\Sigma}_{ik} \mathbf{R}_{kj} = \lambda_j \mathbf{R}_{ij}$$

This is a set of equations that determines the elements of column j in $\mathbf{R}$. This set fits any column in $\mathbf{R}$, and for the sake of convenience I omit the index j from the last expression, so we have:

$$\sum_{k=1}^{n} \boldsymbol{\Sigma}_{ik} \mathbf{R}_{k} = \lambda \mathbf{R}_{i}$$

Now, substituting the elements of $\boldsymbol{\Sigma}$ into the last expression gives:

$$s_i^2 \mathbf{R}_i + \eta_1 \sum_{k=1,k \neq i}^{j} \mathbf{R}_k + \eta \sum_{k=j+1}^{n} \mathbf{R}_k = \lambda \mathbf{R}_i \quad , \quad i = 1, ..., j$$

$$s_i^2 \mathbf{R}_i + \eta \sum_{k=1}^{j} \mathbf{R}_k + \eta_2 \sum_{k=j+1,k \neq i}^{n} \mathbf{R}_k = \lambda \mathbf{R}_i \quad , \quad i = j+1, ..., n$$

Because $\mathbf{R}_i$ exists $\forall i$, we can divide the two expressions by $\dfrac{1}{\lambda - (s_i^2 - \eta_1)}$ and $\dfrac{1}{\lambda - (s_i^2 - \eta_2)}$

respectively. Therefore, after rearranging we obtain:

$$\mathbf{R}_i = \frac{\eta_1}{\lambda - (s_i^2 - \eta_1)} \sum_{k=1}^{j} \mathbf{R}_k + \frac{\eta}{\lambda - (s_i^2 - \eta_1)} \sum_{k=j+1}^{n} \mathbf{R}_k = 0 \quad , \quad i = 1, ..., j$$

$$\mathbf{R}_i = \frac{\eta}{\lambda - (s_i^2 - \eta_2)} \sum_{k=1}^{j} \mathbf{R}_k + \frac{\eta_2}{\lambda - (s_i^2 - \eta_2)} \sum_{k=j+1}^{n} \mathbf{R}_k = 0 \quad , \quad i = j+1, ..., n$$

Summing the two expressions over all the possible values of i (and replacing the index i with k) gives:

$$\sum_{k=1}^{j} \mathbf{R}_k = \eta_1 \sum_{k=1}^{j} \frac{1}{\lambda - (s_k^2 - \eta_1)} \sum_{k=1}^{j} \mathbf{R}_k + \eta \sum_{k=1}^{j} \frac{1}{\lambda - (s_k^2 - \eta_1)} \sum_{k=j+1}^{n} \mathbf{R}_k$$

$$\sum_{k=j+1}^{n} \mathbf{R}_k = \eta \sum_{k=j+1}^{n} \frac{1}{\lambda - (s_k^2 - \eta_2)} \sum_{k=1}^{j} \mathbf{R}_k + \eta_2 \sum_{k=j+1}^{n} \frac{1}{\lambda - (s_k^2 - \eta_2)} \sum_{k=j+1}^{n} \mathbf{R}_k$$

And after rearranging we get:

$$\left[\eta_1 \sum_{k=1}^{j} \frac{1}{\lambda - (s_k^2 - \eta_1)} - 1 \right] \sum_{k=1}^{j} \mathbf{R}_k + \eta \sum_{k=1}^{j} \frac{1}{\lambda - (s_k^2 - \eta_1)} \sum_{k=j+1}^{n} \mathbf{R}_k = 0$$

$$\eta \sum_{k=j+1}^{n} \frac{1}{\lambda - (s_k^2 - \eta_2)} \sum_{k=1}^{j} \mathbf{R}_k + \left[\eta_2 \sum_{k=j+1}^{n} \frac{1}{\lambda - (s_k^2 - \eta_2)} - 1 \right] \sum_{k=j+1}^{n} \mathbf{R}_k = 0$$

These are two homogenous linear equations in two unknowns, $\sum_{k=1}^{j} \mathbf{R}_k$ and $\sum_{k=j+1}^{n} \mathbf{R}_k$. It can be

shown that $\sum_{k=1}^{j} \mathbf{R}_k$ and $\sum_{k=j+1}^{n} \mathbf{R}_k$ cannot both equal zero. A set of homogenous linear equations has a

solution other than the zero solution, if and only if its determinant equals 0. Therefore, and since

$\sum_{k=1}^{j} \mathbf{R}_k$ and $\sum_{k=j+1}^{n} \mathbf{R}_k$ exist, we get:[16]

$$\left[\eta_1 \sum_{k=1}^{j} \frac{1}{\lambda - (s_k^2 - \eta_1)} - 1 \right] \left[\eta_2 \sum_{k=j+1}^{n} \frac{1}{\lambda - (s_k^2 - \eta_2)} - 1 \right] = \eta^2 \sum_{k=1}^{j} \frac{1}{\lambda - (s_k^2 - \eta_1)} \sum_{k=j+1}^{n} \frac{1}{\lambda - (s_k^2 - \eta_2)}$$

Denote:

$$F_1(\lambda) = \sum_{k=1}^{j} \frac{1}{\lambda - (s_k^2 - \eta_1)} \quad \text{and} \quad F_2(\lambda) = \sum_{k=j+1}^{n} \frac{1}{\lambda - (s_k^2 - \eta_2)}$$

Therefore, the last equation becomes:

$$\left[\eta_1 F_1(\lambda) - 1 \right] \left[\eta_2 F_2(\lambda) - 1 \right] = \eta^2 F_1(\lambda) F_2(\lambda)$$

---

[16] $\sum_{k=1}^{j} \mathbf{R}_k$ and $\sum_{k=j+1}^{n} \mathbf{R}_k$ exist because the diagonalizing procedure is well defined here.

Σ is finite. Thus, λ, $s_k^2\ \forall k$, η_1 and η_2 are finite, and therefore $F_i(\lambda)\neq 0$, $i=1,2$. Hence, we can divide the last expression by $F_1(\lambda)F_2(\lambda)$ and obtain:

$$(A13)\quad \left[\eta_1-\frac{1}{F_1(\lambda)}\right]\left[\eta_2-\frac{1}{F_2(\lambda)}\right]=\eta^2$$

It can be shown that for $\lambda\leq 0$:

$$-\infty<\frac{1}{F_i(\lambda)}\leq-\frac{1}{\displaystyle\sum_i\frac{1}{s_k^2-\eta_i}}\quad,\ i=1,2$$

$$\text{where }\Sigma_i=\begin{cases}\displaystyle\sum_{k=1}^{j}&,\ i=1\\[4mm]\displaystyle\sum_{k=j+1}^{n}&,\ i=2\end{cases}$$

And therefore:

$$(A14)\quad \infty>\eta_i-\frac{1}{F_i(\lambda)}\geq\eta_i+\frac{1}{\displaystyle\sum_i\frac{1}{s_k^2-\eta_i}}\quad,\ i=1,2\ ,\ \lambda\leq 0$$

Now, I assume that:

$$(A15)\quad \eta_1>-\frac{1}{\displaystyle\sum_{k=1}^{j}\frac{1}{s_k^2-\eta_1}}\quad,\quad \eta_2>-\frac{1}{\displaystyle\sum_{k=j+1}^{n}\frac{1}{s_k^2-\eta_2}}$$

Thus, from (A14) and (A15) we obtain that:

$$\left[\eta_1-\frac{1}{F_1(\lambda)}\right]\left[\eta_2-\frac{1}{F_2(\lambda)}\right]\geq\left(\eta_1+\frac{1}{\displaystyle\sum_{k=1}^{j}\frac{1}{s_k^2-\eta_1}}\right)\left(\eta_2+\frac{1}{\displaystyle\sum_{k=j+1}^{n}\frac{1}{s_k^2-\eta_2}}\right)\ ,\ \lambda\leq 0$$

27

And if we also assume:

$$(A16) \quad \left(\eta_1 + \cfrac{1}{\sum_{k=1}^{j} \cfrac{1}{s_k^2 - \eta_1}}\right)\left(\eta_2 + \cfrac{1}{\sum_{k=j+1}^{n} \cfrac{1}{s_k^2 - \eta_2}}\right) > \eta^2$$

Then:

$$\left[\eta_1 - \frac{1}{F_1(\lambda)}\right]\left[\eta_2 - \frac{1}{F_2(\lambda)}\right] > \eta^2 \quad , \quad \lambda \le 0$$

which means that under the conditions in (A15) and (A16) there are no nonpositive eigenvalues for which equation (A13) holds. Now, since equation (A13) holds under the conditions in (A15) and (A16), it means that under the conditions in (A15) and (A16) equation (A13) holds only for strictly positive eigenvalues. In other words, the conditions in (A15) and (A16) are sufficient to ensure that Σ is positive definite.[17]

Because of (A15), I can denote:

$$\left|\eta_{12}^*\right| = + \sqrt{\left(\eta_1 + \cfrac{1}{\sum_{k=1}^{j} \cfrac{1}{s_k^2 - \eta_1}}\right)\left(\eta_2 + \cfrac{1}{\sum_{k=j+1}^{n} \cfrac{1}{s_k^2 - \eta_2}}\right)}$$

And to write again the set of the sufficient conditions from (A15) and (A16) as follows:

$$\eta_1 > -\cfrac{1}{\sum_{k=1}^{j} \cfrac{1}{s_k^2 - \eta_1}} \quad , \quad \eta_2 > -\cfrac{1}{\sum_{k=j+1}^{n} \cfrac{1}{s_k^2 - \eta_2}} \quad , \quad -\left|\eta_{12}^*\right| < \eta < \left|\eta_{12}^*\right|$$

---

[17] It can be shown that these sufficient conditions are also necessary. However, I do not show it here, as my goal is to find a set of sufficient conditions on η_1, η_2 and η.

It can be shown that:

$$\eta_1 > -\frac{1}{\sum\limits_{k=1}^{j} \frac{1}{s_k^2 - \eta_1}} \qquad iff \quad -\left|\eta_1^*\right| < \eta_1 \le \min\left(s_k^2\right) \quad, \quad k = 1,...,j$$

$$\eta_2 > -\frac{1}{\sum\limits_{k=j+1}^{n} \frac{1}{s_k^2 - \eta_2}} \qquad iff \quad -\left|\eta_2^*\right| < \eta_2 \le \min\left(s_k^2\right) \quad, \quad k = j+1,...,n$$

where $\left|\eta_1^*\right|$ and $\left|\eta_2^*\right|$ are respectively the unique solutions of the following equations:

$$\left|\eta_1^*\right| = 1 \bigg/ \sum_{k=1}^{j} \frac{1}{s_k^2 + \left|\eta_1^*\right|} \quad \text{and} \quad \left|\eta_2^*\right| = 1 \bigg/ \sum_{k=j+1}^{n} \frac{1}{s_k^2 + \left|\eta_2^*\right|}$$

To sum up, Σ is positive definite if the following set of conditions holds (note that I replace the index k with i):

$$-\left|\eta_1^*\right| < \eta_1 \le \min\left(s_i^2\right) \quad, \quad i = 1,...,j$$

(A17) $\quad -\left|\eta_2^*\right| < \eta_2 \le \min\left(s_i^2\right) \quad, \quad i = j+1,...,n$

$$-\left|\eta_{12}^*\right| < \eta < \left|\eta_{12}^*\right|$$

Now, combining the sets of conditions from (A12) and (A17) gives the sufficient conditions of Proposition 1.

<div align="right">Q.E.D.</div>

The proof of Proposition 2

The first stage of the proof is to derive the weights of the GMVP by directly finding the sums of the rows in the inverted covariance matrix (i.e., without first finding the individual elements of the inverted matrix). For the special case of the two-block matrix, I enclosed a detailed proof for this stage (see phase 1 in the proof of Proposition 1). The same procedure is applied when there are more than two blocks, but due to technical difficulties, it becomes more and more tedious as the number of blocks increases.

Here, with the help of the Mathematica software, I present the general expression for the weight of stock j from block i ($i=1,\ldots,M$; $j=1,\ldots,Z_i$, where Z_i denotes the number of stocks in block i):

$$w_j = \frac{1}{s_j^2 - \eta_i} \cdot \frac{1 + \sum_{k=1;k\neq i}^{M} B_k + \sum_{k,l=1;k<l;k,l\neq i}^{M} B_k B_l + \sum_{k,l,r=1;k<l<r;k,l,r\neq i}^{M} B_k B_l B_r + \ldots + \sum_{k<l<r<\ldots t\ldots<M;k,l,r\ldots,t\ldots,M\neq i} B_k B_l B_r \ldots B_t \ldots B_M}{\sum_{i=1}^{M} A_i C_i}$$

where: $A_i = \sum_{j=1}^{Z_i} \frac{1}{s_j^2 - \eta_i}$, $B_i = (\eta_i - \eta) A_i$, $i = 1,\ldots,M$

and $C_i = 1 + \sum_{k=1;k\neq i}^{M} B_k + \sum_{k,l=1;k<l;k,l\neq i}^{M} B_k B_l + \sum_{k,l,r=1;k<l<r;k,l,r\neq i}^{M} B_k B_l B_r + \ldots + \sum_{k<l<r<\ldots t\ldots<M;k,l,r\ldots,t\ldots,M\neq i} B_k B_l B_r \ldots B_t \ldots B_M$

For example, in a two-block matrix, the weight of stock j from block 1 is:

$$w_j = \frac{1}{s_j^2 - \eta_1} \cdot \frac{1 + B_2}{A_1(1 + B_2) + A_2(1 + B_1)} = \frac{1}{s_j^2 - \eta_1} \cdot \frac{1 + (\eta_2 - \eta) A_2}{A_1 + A_2 + (\eta_1 + \eta_2 - 2\eta) A_1 A_2} \quad , \quad j = 1,\ldots,Z_1 \quad , \text{ as we}$$

had in Proposition 1. In a four-block matrix, for instance, the weight of stock j from block 3 is:

$$w_j = \frac{1}{s_j^2 - \eta_3} \cdot \frac{1 + B_1 + B_2 + B_4 + B_1 B_2 + B_1 B_4 + B_2 B_4 + B_1 B_2 B_4}{\sum_{i=1}^{4} A_i C_i} \quad , \quad j = 1,\ldots,Z_3 .$$

If $0 \leq \eta_i < \text{minvar}_i \ \forall i$, where minvar_i denotes the minimal variance in block i, then

$\frac{1}{s_j^2 - \eta_i}$, $A_i > 0 \ \forall i$ and $\forall j$, $j = 1,\ldots,n$. If also $0 \leq \eta \leq \min(\eta_i)$, then $B_i \geq 0$, $C_i > 0 \ \forall i$. Thus, together these two conditions are sufficient to ensure obtaining a long-only GMVP that is constructed using the general block matrix.

Q.E.D.