JOURNAL OF INVESTMENT MANAGEMENT, Vol. 12, No. 4, (2014), pp. 27–41

www.joim.com

FOR BETTER PERFORMANCE: CONSTRAIN PORTFOLIO WEIGHTS DIFFERENTIALLY AND GLOBALLY

Haim Levy [a] *and Moshe Levy* [b]

Even after more than six decades since the publication of the breakthrough article by Markowitz, the Mean–Variance framework is still the most commonly employed portfolio management tool. Yet, as portfolio managers know all too well, the optimal diversification and the induced performance are very sensitive to potential parameter estimation errors. This paper suggests two new and related portfolio optimization methods to deal with this problem: the Variance-Based Constraints (VBC), and the Global Variance-Based Constraints (GVBC) methods. By the VBC method the constraint imposed on the weight of a given stock is inversely proportional to its standard deviation: the higher a stock's sample standard deviation, the higher the potential estimation error of its parameters, and therefore the tighter the constraint imposed on its weight. GVBC employs a similar idea, but instead of imposing a sharp boundary constraint on each stock, a quadratic "cost" is assigned to deviations from the naive 1/N weight, and a single global constraint is imposed on the total cost of all deviations. We find that these two new methods outperform existing methods. These results are obtained for two different data sets, and are also robust to the number of assets under consideration and to the number of return observations.

1 Introduction

Markowitz's Mean–Variance (MV) rule is the foundation of modern portfolio management. Even after more than six decades, academics and professional investors widely employ the MV framework or some variation of this framework. Yet, employing the MV rule in practice is not a simple task, as the various parameters are unknown. Markowitz himself realizes this fact and he empirically examines the performance of various methods suggested in the literature to deal with the economic loss induced by parameter estimation errors. In this study we suggest

[a]Myles Robinson Professor of Finance, Jerusalem School of Business Administration at the Hebrew University, Jerusalem 91905, Israel. E-mail: mshlevy@huji.ac.il.
[b]Professor of Finance, Jerusalem School of Business Administration at the Hebrew University, Jerusalem 91905, Israel. Tel.: 972 5883219, E-mail: mslm@huji.ac.il.

two novel optimization methods to handle the potential economic loss induced by estimation errors, and show that these methods constitute an improvement relative to existing methods. We hope that the adoption of these new methods will improve the performance of professional portfolio managers.

Deriving the MV efficient portfolios, one implicitly or explicitly assumes that the investors have exact knowledge of expected returns, variances and covariances. In practice, however, the *ex-ante* parameters are unknown and one must rely on estimation, which is generally based on historical returns. The problem is that when naively employing the sample parameters, the MV rule typically yields unacceptable investment strategies, with extreme portfolio weights, and with many assets held short. Moreover, the aggregate percent of the portfolio in short position in the MV optimal investment strategy based on historical returns is generally very large, and may reach hundreds and even thousands of percent. While holding short positions is a legitimate investment strategy, most institutional investors refrain from having large short positions, and typically restrict the investment weights (e.g., see Frost and Savarino, 1988, p. 29).

In addition, one cannot be very confident about the portfolio weights generated by the MV optimizer with sample estimates, because these weights are very sensitive to the sample parameters, especially to the sample means (see Best and Grauer, 1991; Chopra and Ziemba, 1993). This is quite unsettling, as a manager may find that the MV optimal portfolio based on the sample of the last 50 monthly returns, for example, is completely different from the portfolio constructed based on the last 60 months. Numerical analyses of the MV efficiency frontier reveals that even a little reduction in the mean of a given asset may shift the asset from a large long position to a large short position. Thus, even a small difference between the true unknown means and the sample estimated means, quite in the plausible error range, may yield a large deviation between the optimal investment strategy and the one recommended based on the estimated sample parameters.

Because of the key role of the MV rule, numerous studies suggest various ways to deal with the problem of sampling errors. As a consequence of the extreme sensitivity of the optimal diversification weights to the potential errors in the sample estimates of the mean returns, and the typically large errors involved in estimating mean returns, some studies suggest an extreme investment policy which ignores the historical sample means altogether, and focuses on the investment weights corresponding to the sample minimum variance portfolio (see Green and Hollifield, 1992). An even more extreme suggestion is to ignore all historical parameters, means and variance–covariance matrix alike, and to employ the naïve diversification strategy, namely, investing an equal proportion of $1/N$ in each of the N available assets (see, for example, DeMiguel *et al.*, 2009b; Duchin and Levy, 2009). Most other suggested methods incorporate all the sample parameters, but with some modification, either to the parameters or to the optimization procedure. One possibility is to apply "shrinkage" to the sample parameters (Ledoit and Wolf, 2003; Jagannathan and Ma, 2003). Other approaches which consider the sampling errors and suggest methods to mitigate these errors are the Bayesian approach (Jorion, 1986; Markowitz and Usmen, 2003), the Monte Carlo resampling approach (Michaud, 1989, 1998), and the Black–Litterman approach (Black and Litterman, 1992). In the next section we provide a brief review of the main portfolio optimization methods employed.

Which optimization method is better at handling the sampling errors, hence performs best? Is the relative ranking of the various methods invariant to the number of assets under consideration? If constraints are imposed on the portfolio weights to avoid extreme positions due to sampling error, what type of constraints should be imposed, and how stringent should they be? In answering these questions, one needs to consider two potential economic losses that should be balanced in determining the portfolio weights:

(1) Using the historical sample parameters may lead to a portfolio which is very different from the optimal portfolio which is based on the true parameters, due to the sampling errors.
(2) Drastically modifying the sample parameters (or completely ignoring them) or, alternatively, imposing severe constraints on the portfolio weights may lead, once again, to suboptimality and economic loss. For example, if the parameter estimates are very close to the true parameters, imposing stringent constraints on the weights may hamper performance.

Researchers who suggest modifications to the simple MV optimizer are aware of the importance of the historical sample data, as well as the above two possible losses. Hence, most of them do not advocate ignoring the historical data completely, but rather giving it some weight in the portfolio construction procedure. The main issue the investor faces is how much weight to give to the historical information. Indeed, the goal is to find a golden path that balances the two potential losses discussed above. Namely, one needs to employ the information from the sample returns, but simultaneously to restrict extreme diversification policies, which may be due to sampling errors.

In a recent paper, Levy and Levy (2013) suggest two novel portfolio optimization methods, called Variance-Based Constrained optimization (VBC) and Global Variance-Based Constrained optimization (GVBC). These methods are generalizations of the portfolio weight constraints method. The main idea in both new methods is to take into account the fact that the estimation error is not the same for all stocks: the estimation errors are larger for stocks with larger sample variances. The VBC method imposes constraints that are inversely related to the stock's sample standard deviation: the lower the standard deviation, the lower the estimation error, and the looser the constraints imposed on the stock's portfolio weight.

The VBC method, however, does not imply a higher "cost" for more extreme weights, as long as they are within the allowed bounds. It also does not take into account the relative contribution of different stocks to the portfolio's performance. For example, if we have two stocks with the same sample standard deviation but very different sample means, we may want to allow looser constraints for the stock with the higher mean, because it contributes more to the portfolio's mean. The Global VBC (GVBC) method incorporates these two elements by formulating a quadratic "cost" on the deviation of weights from the naïve $1/N$ weight and imposing a single constraint on the total cost of all weights. We explain these two methods in detail in Section 3. We compare these methods with the main other portfolio optimization methods commonly employed in practice. We find that the new methods outperform all other methods, with the GVBC method performing best.

2 The alternative methods examined

This section provides a brief review of the main portfolio optimization methods suggested in the literature. For a more detailed discussion of these methods, we refer the reader to the relevant original papers.

The Mean–Variance Optimizer

Mean–Variance optimization yields the portfolio which maximizes the mean for a given standard deviation, or stated differently, the portfolio with the maximal Sharpe ratio. In the basic application of this method the sample parameters are employed as estimates of the *ex-ante* parameters. Given a set of sample mean *excess* returns, $\hat{\mu}$, and sample covariance matrix, $\hat{\Sigma}$, the MV optimizer is the solution to:

$$\begin{aligned} &\text{Maximize:} \quad \frac{x'\hat{\mu}}{(x'\hat{\Sigma}x)^{1/2}} \\ &\text{s.t.: } x' \cdot \underline{1} = 1, \end{aligned} \tag{1}$$

where x is a vector of N portfolio weights, x' is its transpose, and $\underline{1}$ is a vector of N 1's. Note that as $\hat{\mu}$ denotes the vector of returns in excess of the risk-free rate, and as $x'\hat{\Sigma}x$ is the portfolio variance, the expression in Equation (1) is the portfolio's Sharpe ratio.

This method typically yields extreme portfolio weights. Specifically, as the number of assets under consideration, N, increases, the percentage of assets held short approaches 50%, an unacceptable result by both professional investors and academics who are interested in equilibrium models (for example, see Levy, 1983; Brennan and Lo, 2010; Levy and Ritov, 2011; Levy and Roll, 2014). One possible solution to these many short positions is to directly impose a no-short constraint, which was Markowitz's original suggestion, and his motivation for developing his critical line algorithm (Markowitz, 1959). Maximizing the Sharpe ratio under this constraint typically yields only a small number of assets in the portfolio, where most of the assets are left out (i.e. with portfolio weights of zero).

The Minimum Variance Portfolio (MVP)

Best and Grauer (1991, 1992) show that portfolio weights are very sensitive to the mean return estimates, and Green and Hollifield (1992) argue that the main factor inducing the selection of a portfolio which substantially deviates from the optimal portfolio is the error in the mean return estimates. Therefore, they suggest ignoring the assets' sample means altogether. This implies holding the Minimum Variance Portfolio (MVP), which is given by the solution to:

$$\begin{aligned} &\text{Minimize:} \quad x'\hat{\Sigma}x \\ &\text{s.t.: } x' \cdot \underline{1} = 1. \end{aligned} \tag{2}$$

Unfortunately, the MVP also typically involves extreme portfolio weights and many short positions.

Jagannathan and Ma (2003) add to the MVP approach the no-short constraints, and show that the non-negativity constraints can help in practice. They show that each of the non-negativity constraints is equivalent to reducing the estimated covariance of the asset under consideration with other available assets. As stocks with high positive covariances tend to be held short, any positive estimation error in the covariance estimate may induce suboptimal short position. Thus, imposing the non-negativity constraints reduces the portfolio asset misallocation due to sampling errors. Moreover, they show that imposing upper bounds on the investment proportions reduces the sampling errors and achieve an effect which is similar to the effect induced by parameter shrinkage methods (for more on the shrinkage method, see, for example, Ledoit and Wolf, 2003). Below we present the performance of the MVP approach with and without the short-selling constraints.

Naïve Diversification

The naïve $1/N$ strategy is perhaps the world's first portfolio diversification advice, which is about 1500 years old.[1] While Green and Hollifield suggest ignoring the sample mean returns in seeking the optimal investment weights, the naïve approach suggests ignoring *all* historical information and simply investing $1/N$ in each of the N available assets. Thus, this method implicitly assumes that the sampling errors of the means and variance covariance matrix are very large, hence to protect against possible extreme non-optimal investment policies, it is recommended to diversify equally between the N available assets. In out-of-sample analysis the $1/N$ diversification strategy may have merits when the distributions of returns are very unstable. For the discussion of the performance of the $1/N$ strategy for small and large number of assets see Duchin and Levy (2009), DeMiguel *et al.* (2009a), and Kritzman *et al.* (2010). Tu and Zhou (2011) suggest combining naïve diversification with other optimization strategies by taking the portfolio weights as a weighted average between the naïve weights and the weights suggested by the optimization strategy.

Michaud's Resampling Method

Michaud (1989, 1998) suggests the following procedure in order to avoid large sampling errors. First, estimate the sample means and the sample variance covariance matrix. Then draw T observations from a multivariate normal distribution with these sample parameters, i.e. generate a resampled history of T periods. Next, calculate the various parameters as well as the optimal investment weights based on this resampled history. The weights are constrained to be non-negative in this optimization. Repeat this procedure 500 times, hence obtaining 500 vectors of investment weights. Then, average the portfolio weight for each asset across these 500 vectors. Averaging out 500 times for each portfolio weight attenuates the extreme weights due to sampling errors.

Bayesian Approach

In a nutshell, this method takes into account the fact that the estimation error regarding the mean return induces an increased estimation of the variance (and covariance). For example, suppose that in one sample we obtain a mean return of 10% and a standard deviation of 20%, and that in another sample we obtain a mean return of 15% and the same standard deviation of 20%. Taking into account the fact that the location of the distribution may be centered either about 10% or about 15% provides a distribution with a variance larger than 20%, which has to be taken into account in deriving the optimal investment weights. The Bayesian method takes this factor into account in the parameters' estimation procedure. The Bayesian method can be applied with various priors. Here we employ the diffuse prior as in Markowitz and Usmen (2003). For more details on the Bayesian method see Markowitz and Usmen (2003) and Harvey *et al.* (2008).

Homogeneous Portfolio Weight Constraints

The solution to the MV optimizer as in Equation (1) typically leads to extreme portfolio weights, especially when the number of assets is large. These extreme weights, in turn, imply that the parameter estimation errors usually lead to poor portfolio performance. To mitigate this problem it has been suggested to impose constraints on the portfolio weights (Frost and Savarino, 1988; Jagannathan and Ma, 2003). Namely, one can add the following constraint to optimization problem (1):

$$|x_i| \leq \alpha \quad \text{for all } i, i = 1, 2, \ldots, N$$

where α is a positive constant and N stands for the number of assets under consideration. Alternatively, one can take the naive $1/N$ portfolio as the benchmark and impose the following constraint:

$$\left| x_i - \frac{1}{N} \right| \leq \alpha, \quad \text{for all } i, i = 1, 2, \ldots, N. \quad (3)$$

The formulation in Equation (3) is natural as some specific diversification policies emerge as special cases: the case of $\alpha = 0$ implies the naïve portfolio, which is the benchmark case of "zero information" about the *ex-ante* parameters. The larger α, the looser the constraint, and for $\alpha \to \infty$ we are back to the unconstrained MV optimizer, which is optimal in the other extreme case where perfect information about the *ex-ante* parameters is available.

Each of the above six methods has its pros and cons. While the performance of each of these methods has been previously studied, there is no clear consensus about the relative ranking of their performance. This is for two main reasons: First, most studies compare two or three methods, but not all the six methods simultaneously (which with and without short-selling for the MV optimizer and the minimum variance portfolio increases to eight methods). Second, different studies employ different methodologies and different sample periods to evaluate performance. This paper presents a comparison of all eight methods, as well as the two new methods suggested in the next section, in a unified setting. This comparison provides a bird's-eye perspective on the relative standing of the various optimization methods.

3 Two new methods: VBC and GVBC

In these methods the Sharpe ratio (1) is maximized under constraints on the portfolio weights. However, in contrast with the homogeneous constraints method, the idea in both new methods is to impose more stringent constraints on stocks with relatively high standard deviations, as the estimation errors for these stocks' parameters, and hence the potential economic loss, are larger than for stocks with relatively low standard deviations. Below we describe the two suggested new methods, and then provide the intuition and motivation for them with a numerical example.

Variance-Based Constraints (*VBC*)

In the VBC method we suggest replacing the homogeneous constraint $|x_i - \frac{1}{N}| \leq \alpha$ with:

$$\left| x_i - \frac{1}{N} \right| \frac{\sigma_i}{\bar{\sigma}} \leq \alpha, \quad \text{for all } i, i = 1, 2, \ldots, N, \tag{4}$$

where $\bar{\sigma} \equiv \frac{1}{N} \sum_{i=1}^{N} \sigma_i$ is the average standard deviation of all stocks. Equation (4) implies ***differential*** constraints: the higher σ_i (relative to the average standard deviation) the tighter the constraint imposed on the weight of stock i. Note that if all stocks have the same standard deviation, the VBC method reduces to the homogeneous constraint method, as in Equation (3). One should also note that imposing the variance-based constraints is not intended to reduce portfolio variance, but rather, to reflect the larger potential sampling errors for high-variance stocks.

Global Variance-Based Constraint (*GVBC*)

The VBC method suggests that the higher the asset's variance, the tighter the constraint imposed on its weight. Thus, it implies identical constraints for all stocks with the same standard deviation. Recalling that the goal is to find a set of constraints which maximizes the *ex-ante* Sharpe ratio, this may be suboptimal, because the contribution of different stocks with the same standard deviation to the portfolio's Sharpe ratio may be different. In addition, the VBC method does not attach any "cost" to deviations from the benchmark naïve weights, as long as these deviations are within the allowed bounds. The GVBC method addresses these two issues by imposing a single *global* constraint (rather than a specific constraint on each stock) on the vector of portfolio weights:

$$\sum_{i=1}^{N} \left(x_i - \frac{1}{N} \right)^2 \frac{\sigma_i}{\bar{\sigma}} \leq \alpha. \tag{5}$$

Note that the higher the asset's standard deviation, the more costly are deviations of its weight from $1/N$. Therefore the GVBC optimization will tend to restrict extreme positions for such stocks. This is similar to the VBC method. However, in contrast to the VBC method, we may allow different deviations for stocks with the same standard deviations. For example, suppose that we have two stocks with the same standard deviations, but one with a higher sample mean than the other. The GVBC method will allow a larger deviation for the stock with the higher mean, and thus will assign a higher weight for this stock. Thus, we have ***global*** constraint (α) and each stock will get a different share of this α, depending on its standard deviation as well as its contribution to the Sharpe ratio.

Numerical Example

The following example may help explain the logic behind the VBC and GVBC methods, and contrast them with one another and each one of them with the standard homogeneous constraints method. Consider the six stocks with sample parameters given in Table 1. Assume for simplicity of the presentation that the sample correlations are zero. Column (3) in the table shows the portfolio weights obtained by maximizing the portfolio's Sharpe ratio with the standard unconstrained MV optimization (Equation (1)). Column (4) shows the equal weights of the naïve portfolio, and column (5) depicts the weights with the standard homogeneous constraints method, as given in Equation (3), with $\alpha = 2\%$. As column (3) shows, the unconstrained optimization implies that the weights of the first three stocks are much lower than the naïve weight of 16.67%, while the weights of the last three stocks are larger than the naïve weight. Also, note that in this case the first three stocks have the same weight as they have the same parameters. The simple homogenous constraint with $\alpha = 2\%$ allows weights to deviate by no more than 2% from 16.67% (see Equation (3)), which leads to the first three stocks' weights being constrained at $16.67\% - 2\% = 14.67\%$, and the last three stocks' weights at $16.67\% + 2\% = 18.67\%$. Recalling that the goal is the maximization of the Sharpe

Table 1 Numerical illustration of the different constrained optimization methods.

	(1) Sample mean (%)	(2) Sample standard deviation (%)	(3) MV optimizer weights (%)	(4) Naïve weights (%)	(5) Homogeneous constraints weights (%)	(6) VBC weights (%)	(7) GVBC weights (%)
Stock 1	5	8	9.64	16.67	14.67	14.85	11.01
Stock 2	5	8	9.64	16.67	14.67	14.85	11.01
Stock 3	5	8	9.64	16.67	14.67	14.85	11.01
Stock 4	12	8	23.13	16.67	18.67	19.00	21.90
Stock 5	26	12	22.27	16.67	18.67	18.22	21.06
Stock 6	30	12	25.70	16.67	18.67	18.22	24.02

Stock return parameters and portfolio weights are given in percent. $\alpha = 2\%$. The homogeneous constraints method imposes the same constraints to all stocks. The VBC method imposes tighter constraints for stocks with higher standard deviations (compare stocks 4 and 5). The GVBC method assigns a quadratic cost to deviations from the naïve weight, and imposes a single constraint on the total cost. This allows more flexibility to assign higher portfolio weights to stocks that contribute more to the portfolio's performance (compare stocks 5 and 6).

ratio, if the constraints are not binding it is possible that the stock with the higher mean will have a higher weight. However, if the constraint is binding, as in our example, the stock's mean, namely its contribution to Sharpe ratio, does not affect the weight optimal selection.

The homogeneous constraints method, however, does not take into account the fact that the stock's standard deviations are different—which means that we have higher confidence about the parameters of the first four stocks (with $\sigma = 8\%$) than we do about the parameters of stocks 5 and 6 (with $\sigma = 12\%$). As the larger the standard deviation the larger the potential economic damage induced by sampling errors corresponding to the mean return estimates, we suggest the VBC method. Indeed, the VBC method (Equation (4)) takes the different standard deviations into account, as revealed by the weights in column (6). If no constraints were imposed, the weights of stocks 4–6 would all have been higher, and the weights of stocks 1–3 would have been lower (see column 3). The weights under VBC shown in column (6) reveal, of course, perfect symmetry between the first three stocks. Notice, however, that the VBC method assigns a higher weight to stock 4 than to stocks 5 and 6, because of its lower standard deviation.[2]

Finally, like in the homogeneous constraints method, the stock mean return may affect the weight only if the constraint is not binding. A comparison of the weight of stocks 5 and 6 reveals that the same weight is assigned to these two stocks despite the fact that both have the same standard deviation and stock 6 has a higher mean. This occurs because the constraint on the weight given by Equation (3) is binding.

We believe that taking account of the heterogeneous standard deviations offers an improvement relative to the homogeneous constraints method. However, this method still misses an important point: the fact that the relative contribution of different stocks to the portfolio performance (Sharpe ratio) is not the same. Specifically, the mean returns of the various assets are completely ignored in case where the imposed constraints are binding. For instance, the VBC method assigns the same weight of 18.22% to stocks 5 and 6, even though stock 6 has exactly the same sample standard deviation as stock 5, but a higher sample mean. Common sense would suggest assigning a higher weight to stock 6 than to stock 5 in this case. Indeed, this is exactly what the GVBC method yields, as shown in column (7). By the GVBC method deviations from the naïve portfolio weights have a quadratic cost, and the only constraint is on the total cost. This is opposed to the homogeneous constraints and VBC methods, that impose a separate boundary constraint for each stock. Thus, the GVBC method gives more flexibility to allow larger deviations for stocks that contribute more to the portfolio performance. In addition, the continuous cost function seems to be more reasonable than the binary boundary constraint—it assigns a cost to any deviation, rather than being indifferent to any deviations within the allowed limits, as in the homogeneous and VBC methods.

4 Methodology

We conduct a "horse race" between the various portfolio optimization strategies described in the previous two sections. Following the framework in Frost and Savarino (1988), Michaud (1989), Markowitz and Usmen (2003), and Harvey *et al.* (2008), we take N assets with true mean excess returns μ and covariance matrix Σ, which are known to a "referee", but not to investors employing the different strategies which are based on sample parameters. Thus, investors construct their portfolios based on a sample they observe from the true distribution, and the referee evaluates the performance of these portfolios based on the true parameters. In order for the true parameters to be realistic, we take μ and Σ as the sample estimates of actual assets: one set of assets we use is the set stocks of 100 largest U.S. firms (according to December 2011 market values) with complete 20-year monthly return records, and the other set is composed of the 48 Fama–French industry portfolios. Parameters are estimated for monthly returns, in the period 1992–2011 for the first set, and the period 1969–2011 for the second set.

The multivariate distribution is assumed to be normal and the parameters given by μ and Σ are considered as the true parameters, and the referee evaluates the *ex-ante* performance with respect to these parameters. The investors do not know the true parameters. They observe a sample of T returns for each asset, drawn from a multivariate normal distribution with the true parameters. Thus, they observe sample parameters which are generally different from the true parameters. Given these sample parameters, each strategy S yields a set of portfolio weights, x_s, as explained in the previous sections. These weights, in turn, determine the true *ex-ante* portfolio performance for the strategy, calculated with the true parameters: $\mu_P = x_s'\mu, \sigma_P^2 = x_s'\Sigma x_s$.

We repeat this procedure for 100 different samples, drawn from the same true multivariate distribution. The performance of each strategy is evaluated across all 100 scenarios. Namely, for each sample we have a different normal distribution of portfolio returns. We calculate the portfolio mean return and standard deviation for each sample and then we construct the return distribution which is a mixture of these 100 distributions, each with weight 1/100. The Sharpe ratio is (numerically) calculated for this "grand" mixed distribution (for more detail see Levy and Levy, 2013).

5 Results

Ten optimization methods are compared: the MV optimizer, the MV optimizer with no short positions, the minimum variance portfolio, the minimum variance portfolio with no shorts, naïve optimization, Resampling, Bayes, homogeneous constraints, and the two new methods: the Variance-Based Constraints (VBC), and Global Variance-Based Constraints (GVBC). The results of the last three methods, based on portfolio weight constraints, depend on the value of the constraint imposed, α. Figure 1 shows the Sharpe ratios of the various strategies as a function of α, where the assets are the 100 largest U.S. firms. Obviously, the results for the first seven strategies, which do not depend on α, appear as horizontal lines in the figure.

Looking first at the investment strategies which are independent of the constraint, α, we find that the best strategy out of these seven strategies is

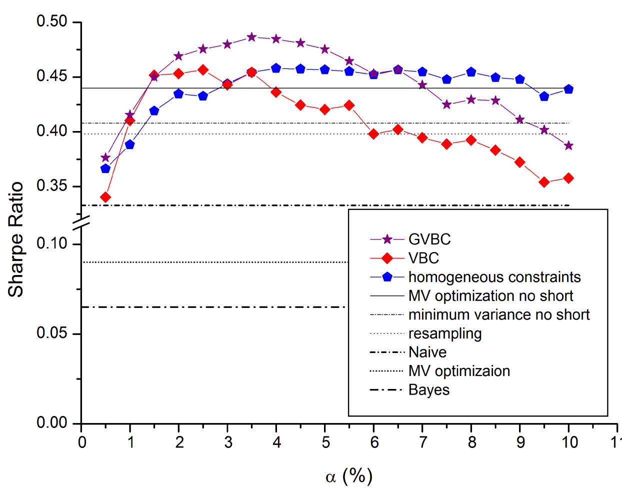

Figure 1 Sharpe ratios of the various strategies. The three constraints strategies as a function of α. Strategies that do not depend on α appear as horizontal lines. Note that the y-axis is broken, to show the MV optimizer and Bayes strategies. The minimum variance strategy yields an intermediate Sharpe ratio of 0.229 (see Table 2), and is not shown in the figure because of the break in the y-axis.

the MV optimizer with no short positions, followed by the minimum variance portfolio with no short positions, the Resampling method, the naïve optimization, and the minimum variance portfolio with short (not shown in the figure because of the break in the y-axis, but see the performance ranking in Table 2). The worst strategies are the MV optimizer (with short) and Bayes, which are very close.[3] The result revealing that the Resampling method outperforms the Bayes strategy is consistent with the results reported by Markowitz and Usmen (2003).

Now let us turn to the performance of the three constraints strategies, which are function of the size of the constraint, α. For very small α all three strategies converge to the naïve strategy, as the portfolio weights converge to $1/N$. For very large α, the strategies converge to the unconstrained MV optimizer, as the constraints become ineffective (this is not evident in the figure, because this convergence occurs at very high values of α). The most important result is that there is a range of α where all three constraints methods outperform all other methods.

Two comments regarding the choice of α are called for. First, note that the investor does not know in advance neither the optimal value of α, nor the corresponding Sharpe ratio, which are known only to the referee. Therefore, the selected value of α, which is at the investor's discretion, may be non-optimal. However, the precise purpose of this exercise is to calculate the Sharpe ratio for various values α, hence the optimal value is revealed at the end of the exercise by the referee, and this knowledge can be adopted in practice. Those investors who mistakenly considered employing a non-optimal α value before seeing the results of this study may adjust their α according to the results of this paper. Secondly, one should not expect a single value of α to be optimal in all cases. The optimal value will generally depend on the nature of the assets under consideration. For example, the higher the variance of the assets' returns, the smaller the optimal value α one would expect. Therefore, when a portfolio of portfolios (e.g., a portfolio of ETFs) is considered, α would generally be relatively large, as the variance of the ETFs is typically small relative to the variance of individual assets. The optimal value of α may also depend on the number of assets included in the portfolio. This is particularly true for the GVBC method where a single constraint is imposed on all assets included in the portfolio. Hence, each strategy should be evaluated at the value of α that maximizes its Sharpe ratio. Although we present the Sharpe ratio for various values of α, only the value which maximizes the Sharpe ratio is relevant to the investor. Figure 1 reveals that the best performance is obtained by the GVBC method with $\alpha = 4\%$.

Table 2 reports the Sharpe ratios of the 10 strategies. In order to examine the dependence of the results on the number of assets, we repeat the analysis for 20 stocks and five stocks, randomly drawn from the set of 100 stocks. In general, the Sharpe ratio increases with the number of assets, because of the greater benefit of diversification. As the table reveals, the three constraints methods yield the highest Sharpe ratios in all three cases, and the GVBC method always ranks first. The differences between the various optimization methods are most pronounced when the number of assets is large. In the case of only five assets the differences across strategies are rather small, except for the unconstrained MV optimizer and Bayes methods, which are always significantly worse than the rest.

Table 3 reports the short positions for each of the six strategies which allow short-selling (the naïve, resampling, MV, and mean–variance optimization with no short-selling strategies are obviously absent in Table 3). This information is important, as some investors would reject any investment

Table 2 Performance of the 10 different strategies with 5, 20, and 100 assets.

	5 Stocks		20 Stocks		100 Stocks	
Method	Sharpe ratio	Rank	Sharpe ratio	Rank	Sharpe ratio	Rank
GVBC	0.269	1	0.330	1	0.486	1
VBC	0.267	2	0.328	2	0.456	3
Homogeneous constraints	0.265	3	0.327	3	0.457	2
Naïve ($1/N$)	0.262	4	0.305	8	0.333	7
Resampling	0.261	5	0.320	4	0.398	6
Minimum variance no short	0.259	6	0.317	5	0.408	5
MV optimization no short	0.245	8	0.311	7	0.440	4
Minimum variance	0.257	7	0.315	6	0.229	8
MV optimization	0.149	9	0.224	10	0.090	9
Bayes	0.131	10	0.228	9	0.065	10

The number of periods in the sample observed by the investor is $T = 120$. The constraints methods provide the highest Sharpe ratio, and the GVBC method is best, in all three cases. The differences between the various methods are most pronounced when the number of assets is large.

Table 3 Short positions as implied by the different methods.

	Proportion of stocks held short (%)			Total short position		
Method	5 Stocks	20 Stocks	100 Stocks	5 Stocks	20 Stocks	100 Stocks
GVBC	0	16.7	38.3	0	0.08	0.79
VBC	0	24.5	47.8	0	0.04	0.90
Homogeneous constraints	0	0	44.7	0	0	0.81
Minimum variance	11.8	29.9	46.8	0.02	0.30	4.34
Bayes	20.0	43.1	48.7	0.26	2.67	8.93
MV optimization	27.4	40.9	47.5	0.37	1.56	13.10

The left-hand side of the table shows the percentage of the *number* of assets held short, while the right-hand side shows the *total short position*. For example, the MV optimization method with 100 stocks implies that for each \$1 invested, \$13.1 are held short and \$14.1 are held long.

strategies with large short positions. Moreover, the large short position may indicate a position which is due to sampling errors, hence may shed some light on the performance reported in Table 2 and on the importance of the imposed constraints on the investment weights.

The left-hand side of the table reports the average percentage of assets in short position. (Recall that we have 100 samples drawn, and each strategy yields a portfolio in each one of these cases). For example, with 100 assets and the MV optimization, on average 47.5 of the assets (which in the 100 asset case is also equal to 47.5%) are held in short positions, and 52.5 of the assets are held long. Two phenomena emerge from this part of the table: the percentage of assets in short positions increases with the number of assets under consideration, and for the 100 asset case all methods yield a percentage of 38.3%–48.7%

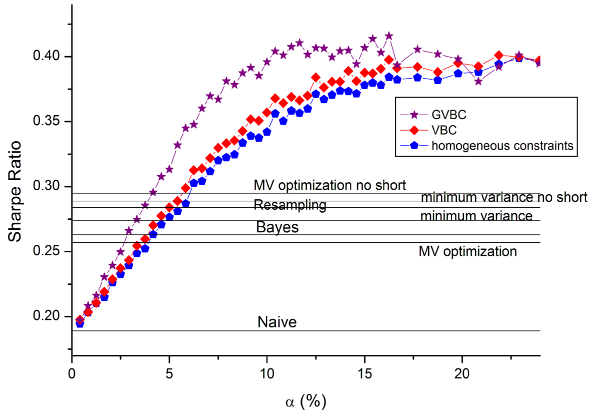

Figure 2 Sharpe ratios of the various strategies for the 48 Fama–French industry portfolios. The best strategies are GVBC, followed by VBC and homogeneous constraints. Notice that relative to the case of 100 individual stocks the optimal α levels are higher. This is most likely due to the fact that the industry portfolios are much less volatile than the individual stocks, and the sampling errors are therefore smaller.

of the assets held short. This is consistent with the results of Levy (1983) and Levy and Ritov (2011). Thus, considering short positions as a deficiency, it seems that for large portfolios all six strategies suffer from this deficiency.

However, the right-hand side of the table is more meaningful for examining the short position effect, as it reports the percentage of the *investment* in short positions rather than the percentage of the *number* of assets in short position. Here we obtain large differences among the various methods. For example, with the MV optimization, for each \$1 investment we have \$13.10 in short position (and therefore \$14.10 in long position). Thus, the short position in this case is 1,310%. These extreme weights explain the low Sharpe ratios of the MV optimizer, Bayes method, and the minimum variance portfolio. With the three constraints methods the short positions are much smaller, and are less than 100%. Short positions are lowest in the GVBC method.

In order to examine how the different strategies compare when the assets are themselves portfolios, rather than individual stocks, we apply the analysis to the set of 48 Fama–French industry portfolios. The results are shown in Figure 2 and Table 4. Figure 2 shows the Sharpe ratios when the sample observed by the investor consists of $T = 120$ periods, (the same as in the preceding

Table 4 Performance of the various strategies with the 48 Fama–French industry portfolios.

	30 Return observations		60 Return observations		120 Return observations	
Method	Sharpe	Rank	Sharpe	Rank	Sharpe	Rank
GVBC	0.310	1	0.353	1	0.416	1
VBC	0.309	2	0.351	2	0.401	2
Homogeneous constraints	0.298	3	0.340	3	0.399	3
Resampling	0.265	5	0.284	4	0.284	6
Minimum variance no short	0.267	4	0.282	5	0.289	5
MV optimization no short	0.239	6	0.265	6	0.295	4
Naïve	0.189	7	0.191	7	0.189	10
Minimum variance	−0.006	8	0.159	8	0.274	7
Bayes	−0.009	9–10	0.007	9	0.257	9
MV optimization	−0.009	9–10	−0.001	10	0.263	8

The investors observe a sample of $T = 30/60/120$ returns drawn from the true return distribution. In all three cases GVBC is ranked first, VBC second, and the homogeneous constraints method is ranked third.

100-stock case). In Table 4 we also report the performance with $T = 60$ and $T = 30$ observations. Obviously, the lower T, the larger the estimation errors and the lower the Sharpe ratio for all strategies, except for the naïve strategy which does not dependent on the sample at all. It is interesting that the ranking is relatively stable across the different values of T: GVBC is first, VBC is second, and the homogeneous constraints method is third in all cases. This is very similar to the results obtained for the 100 stocks. Figure 2 shows that performance is best for values of α that are higher than in the case of individual stocks. The optimal value is in the range 15%–20%. This is most likely due to the fact that the industry portfolios are less volatile than individual stocks: the average industry portfolio monthly standard deviation is 5.5%, compared to an average of 8.8% for individual stocks. This lower volatility implies lower estimation error, and thus one does better with looser constraints in this case.

6 Conclusion

The MV rule is the most commonly employed portfolio decision rule, in both academic research and in practical investing. A central problem in the practical implementation of this rule is that of parameter estimation errors, and therefore many methods have been developed to deal with these errors. Levy and Levy (2013) suggest two new portfolio optimization methods that are extensions of the homogeneous constrained optimization method: the Variance-Based Constraint (VBC) method and the Global Variance-Based Constraint (GVBC) method. The idea behind the VBC method is straightforward: for assets with relatively high variances the parameter estimation errors are relatively large, hence the potential economic loss induced by sampling errors is relatively large. Therefore, the higher the asset's variance, the tighter the constraint imposed on the asset's weight. Put differently, we have more confidence in the parameters of assets with lower sample variances, and therefore we allow the weights of these assets to deviate more from the "zero information" benchmark of the naïve $1/N$ weight.

The advantage of the GVBC method relative to VBC is that rather than imposing a sharp boundary constraint on the weight of each asset, it assigns a quadratic cost to any deviation from the naïve weight, depending again on the variance, and it imposes a single global constraint on the total cost for all assets. This allows imposing looser constraints on assets that contribute more to the Sharpe ratio. It is important to emphasize that in the homogeneous constraints method and in the VBC method the mean returns of the various assets play no role in the optimization if the constraints are binding. In contrast, in the GVBC method, as there is a global constraint rather than a specific constraint on each asset, the higher the mean of an asset, other things being kept equal, the higher the weight of the asset. This gives an edge to the GVBC method over the other methods.

In this paper we employ the cornerstone mean–variance framework, and thus we take the investor's objective function as the maximization of her portfolio's Sharpe ratio. However, the same logic of constraining an asset's portfolio weight based on the asset's variance is general and applies to other frameworks as well. For example, in the four-factor framework the investor's objective function may be the maximization of the portfolio's excess return α relative to the portfolio's factor exposures. In this framework, determining the portfolio weights depends on the estimation of the individual asset parameters, which are again subject to estimation error. Thus, here too, the investor can benefit by imposing constraints on portfolio weights. The main point of this paper is that one would want to impose the most stringent constraints on the assets with the highest volatilities, for which the sampling errors are the largest.

We compare the performance of 10 commonly employed optimization methods: the unconstrained MV optimizer, no-short MV optimizer, minimum variance portfolio, no-short minimum variance portfolio, naïve optimization, resampling, Bayes method, homogeneous constraints on portfolio weights, VBC, and GVBC. We find that the constrained optimization methods outperform the other methods. The best results are obtained with the GVBC method. These results are robust to the length of the observed sample period, and the set of assets considered. The differences between the various methods are most pronounced when the number of assets is large.

It is well known that imposing no-short constraints typically improves performance. This is certainly true for the MV optimizer and the minimum variance portfolio, methods that yield very extreme short positions if they are unconstrained. However, we find that the VBC and GVBC methods work best without imposing the additional no-short constraint. These methods yield relatively small short positions, and these moderate positions enhance portfolio performance. In situations in which the no-short constraint is legally imposed, the VBC and GVBC methods can be employed with this additional constraint.

Frost and Savarino's (1988) classic title is: "For Better Performance: Constrain Portfolio Weights". The main message of this study is an addition of three words to this title: For Better Performance: Constrain Portfolio Weights ***Differentially and Globally***. As the application of our suggested methods is simple to implement, we hope they will be adopted by professional investors, and lead to an improvement in investment performance.

Notes

[1] The Babylonian Talmud recommends: "Man should always divide his wealth into three parts: one-third in land, one-third in commerce and one-third retained in his own hands" (Baba Metzia 42a), see also Levy and Sarnat (1972, p. 3).

[2] The average standard deviation of the six assets given in Table 1 is $\bar{\sigma} = 9.33\%$, thus the constraint on the first four stocks, with $\sigma = 8\%$, is $|x_i - \frac{1}{N}| \le \alpha \frac{\bar{\sigma}}{\sigma_i} = 2\frac{9.33}{8} = 2.33\%$. The constraint on the last two stocks, with $\sigma = 12\%$, is $|x_i - \frac{1}{N}| \le \alpha \frac{\bar{\sigma}}{\sigma_i} = 2\frac{9.33}{12} = 1.55\%$. Accordingly, the weights of the last two stocks are $16.67\% + 1.55\% = 18.22\%$, and the weight of stock 4 is $16.67\% + 2.33\% = 19\%$. Note that in addition to the individual constraints on each of the stocks, we also have the constraint $\sum_{i=1}^{6} x_i = 1$. This means that the total positive deviations from $1/N$ must be balanced with the total negative deviations. In this example, the direct constraints on the positive deviations (stocks 4–6) are more stringent than the constraints on the negative deviations (stocks 1–3), because of the higher standard deviations of stocks 5 and 6. Thus, the effective constraint on the first three stocks is not $|x - 16.67| \le 2.33$, but rather the combination of the constraint $\sum_{i=1}^{6} x_i = 1$ plus the constraints on stocks 4–6.

[3] In fact, we find that the Bayes method always yields results very close to that of the MV optimizer. This can be interpreted in the following way. The Bayes method adds to the sample variance a correction term which is the variance of the sample mean across samples (and similarly for covariances, see Equations (7)–(9) in Markowitz and Usmen, 2003). The variance of the sample mean is σ^2/T, where T is the sample size. Thus, the correction to the variance is of order $1/T$, which is typically very small. For example, with $T = 100$ observations the correction to the sample variance is only 1% of the sample value. Hence, the Bayes weights are very similar to the MV optimizer weights.

References

Best, M. J. and Grauer, R. (1991). "On the Sensitivity of Mean–Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results," *Review of Financial Studies* **4**, 315–342.

Best, M. J. and Grauer, R. (1992). "Positively Weighted Minimum Variance Portfolios and the Structure of Asset

Expected Returns," *Journal of Financial and Quantitative Analysis* **27**, 4.

Black, F. and Litterman, R. (1992). "Global Portfolio Optimization," *Financial Analysts Journal*, 28–43.

Brennan, T. J. and Lo, A. W. (2010). "Impossible Frontiers," *Management Science* **56**, 905–923.

Chopra, V. K. and Ziemba, W. T. (1993). "The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice," *Journal of Portfolio Management* **19**, 6–11.

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

DeMiguel, V., Garlappi, L., Nogales, F. J., and Uppal, R. (2009b). "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms," *Management Science* **55**, 798–812.

Duchin, R. and Levy, H. (2009). "Markowitz versus the Talmudic Portfolio Diversification Strategies," *Journal of Portfolio Management* **35**, 71–74.

Frost, P. A. and Savarino, J. E. (1988). "For Better Performance: Constrain Portfolio Weights," *Journal of Portfolio Management* **15**, 29–34.

Green, R. C. and Hollifield, B. (1992). "When Will Mean/Variance Efficient Portfolios be Well Diversified?" *Journal of Finance* **47**, 1785–1809.

Harvey, C. R., Liechty, J. C., and Liechty, M. W. (2008). "Bayes vs. Resampling: A Rematch," *Journal of Investment Management* **6**, 1–17.

Jagannathan, R. and Ma, T. (2003). "Risk Reduction in Large Portfolios: A Role for Portfolio Weight Constraints," *Journal of Finance* **58**, 1651–1684.

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

Kritzman, M., Page, S., and Turkington, D. (2010). "In Defense of Optimization: The Fallacy of $1/N$," *Financial Analysts Journal* **66**, 31–39.

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

Levy, H. (1983). "The Capital Asset Pricing Model: Theory and Empiricism," *Economic Journal* **93**, 145–165.

Levy H. and Levy, M. (2004). "Prospect Theory and Mean–Variance Analysis," *Review of Financial Studies* **17**, 1015–1041.

Levy H. and Levy, M. (2013). "The Benefits of Differential Variance-Based Constraints in Portfolio Optimization," *European Journal of Operational Research*, forthcoming.

Levy, H. and Markowitz, H. M. (1979). "Approximating Expected Utility by a Function of Mean and Variance," *American Economic Review* **69**, 308–317.

Levy, H. and Sarnat, M. (1972). *Investment and Portfolio Analysis*. New York: Wiley.

Levy, M. and Ritov, Y. (2011). "Mean–Variance Efficient Portfolios with Many Assets: 50% Short," *Quantitative Finance* **11**, 1461–1471.

Levy, M. and Roll, R. (2014). "Possible and Impossible Frontiers," *Critical Finance Review*, forthcoming.

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

Markowitz, H. M. (1959). *Portfolio Selection*: *Efficient Diversification of Investments*. New York: John Wiley & Sons.

Markowitz, H. M. (1991). "Foundations of Portfolio Theory," *Journal of Finance* **46**, 469–477.

Markowitz, H. M. (1993). "Computation of Mean-Semivariance Efficient Sets by the Critical Line Algorithm," *Annals of Operational Research* **45**, 307–317.

Markowitz, H. M. and Usmen, N. (2003). "Resampled Frontiers versus Diffuse Bayes: An Experiment," *Journal of Investment Management* **1**, 9–25.

Michaud, R. O. (1989). "The Markowitz Optimization Enigma: Is 'optimized' optimal?" *Financial Analysts Journal*, 31–42.

Michaud, R. O. (1998). *Efficient Asset Management*. Boston: Harvard Business School Press.

Tu, J. and Zhou, G. (2011). "Markowitz Meets Talmud: A Combination of Sophisticated and Naïve Diversification Strategies," *Journal of Financial Economics* **99**, 204–215.

Keywords: Mean–variance analysis; portfolio optimization; portfolio constraints; estimation error