European Journal of Operational Research 253 (2016) 418–427

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

Value of information in portfolio selection, with a Taiwan stock market application illustration

Chiang Kao[a], Ralph E. Steuer[b,*]

[a] Department of Industrial and Information Management, National Cheng Kung University, Tainan 70101, Taiwan, ROC
[b] Department of Finance, University of Georgia, Athens, GA 30602-6253, USA

ARTICLE INFO

Article history:
Received 2 October 2015
Accepted 6 February 2016
Available online 11 February 2016

Keywords:
Efficient points
Portfolio selection
Value of information
Piecewise linear paths
Parametric quadratic programming

ABSTRACT

Despite many proposed alternatives, the predominant model in portfolio selection is still mean–variance. However, the main weakness of the mean–variance model is in the specification of the expected returns of the individual securities involved. If this process is not accurate, the allocations of capital to the different securities will in almost all certainty be incorrect. If, however, this process can be made accurate, then correct allocations can be made, and the additional expected return following from this is the value of information. This paper thus proposes a methodology to calculate the value of information. A related idea of a level of disappointment is also shown. How value of information calculations can be important in helping a mutual fund settle on how much to set aside for research is discussed in reference to a Taiwan Stock Exchange illustrative application in which the value of information appears to be substantial. Heavy use is made of parametric quadratic programming to keep computation times down for the methodology.

1. Introduction

The problem of portfolio selection – on how to invest a sum of money across a series of assets for optimal return – continues to be a challenge, as it always has been ever since there has been accumulated wealth in the world. Let there be a beginning of a holding period and an end of the holding period. Also, let r_i be the return on asset i over the holding period, and w_i be the proportion of initial capital invested in asset i at the beginning of the holding period, and held in asset i throughout the holding period. With the goal being to maximize end of period wealth, and with constraints in canonical form, the problem of portfolio selection is

$$\begin{aligned} &\max\ r_p = \sum_{i=1}^{n} r_i w_i \\ &s.t. \sum_{i=1}^{n} w_i = 1 \\ &w_i \geq 0 \quad \text{for all } i \end{aligned} \tag{1}$$

where r_p is the return on one's capital over the holding period, n is the number of securities eligible for inclusion in a portfolio, and the sum-to-one constraint along with the nonnegativity restrictions define the feasible region in decision space. In the model, r_p is portfolio return, and with the w_i weights arranged in the form of $\mathbf{w} = (w_1, \ldots, w_n)$, $\mathbf{w}$ is called a fund allocation vector.

The problem looks innocuous enough, as it appears to be a linear programming problem, which in fact it is, except for the objective function. The difficulty in the objective function is that the r_i, the returns of the individual securities over the holding period, are random variables, and hence r_p is a random variable. Portfolio selection is thus the problem of maximizing the random variable of portfolio return – but to do so it is necessary to make decisions on the w_i at the beginning of the holding period based upon the values of the r_i that are not known until the end of the holding period. This makes Model (1) a stochastic programming problem. In this form, the problem of portfolio selection has been much discussed and analyzed. Thousands of papers have been written on the problem as the basic model can take on many related forms.

As defined by Caballero, Cerdá, Muñoz, Rey, and Stancu-Minasian (2001), if in a programming problem some of the parameters take unknown values at the time of making a decision, and these parameters are random variables, then the problem is a stochastic programming problem. Stochastic programming problems are notoriously difficult to solve, and solution methods are usually developed based upon the type of problem being considered (Beraldi, Violi, & Simone, 2011; Shapiro & Philpott, 2007).

* Corresponding author. Tel.: +1 7065423782.
E-mail address: rsteuer@uga.edu (R.E. Steuer).

http://dx.doi.org/10.1016/j.ejor.2016.02.011

One often contemplated approach is to employ interpretations and assumptions so as to strive for an equivalent deterministic formulation that can be solved in a reasonably straightforward fashion.

Through reasoning such as overviewed in many places as in Huang and Litzenberger (1988), it is generally accepted that investors are expected utility maximizers. Under this assumption, and where U is the investor's utility function, Model (1) can be rewritten equivalently as

$$\begin{aligned} &\max\ E[U(r_p)] \\ &s.t. \sum_{i=1}^{n} w_i = 1 \\ &w_i \geq 0 \quad \text{for all } i \end{aligned} \tag{2}$$

An advantage of Model (2) is that all random variables have been cleared from the formulation. With investors assumed to possess declining marginal utility, U is at least known to be concave and increasing.

Two schools of thought have evolved on how to address Model (2) with its expected utility objective function. One is to try to acquire enough information about the decision maker's preferences to enable the creation of an optimization problem that can be solved directly for an optimal portfolio. The "safety first" strategy of Roy (1952) is an example of this approach. More recent examples involving a range of techniques, although in the multi-criteria arena, can be found, for instance, in Ballestero and Romero (1996), Arenas Parra, Bilbao Terol, and Rodríguez Uría (2001), Bilbao-Terol, Pérez-Gladish, Arenas-Parra, and Rodríguez-Uría (2006), Abdelaziz, Aouni, and El-Fayedh (2007), Fang, Lai, and Wang (2008), and Aouni, Colapinto, and La Torre (2014). But these techniques are difficult because the setting up of the optimization problem generally requires more knowledge about the optimal solution to be found than is possible beforehand. The other school of thought involves parameterizing U and then attempting to solve Model (2) for all unknown values of U's parameter(s).

Now, if U is quadratic, which is a common assumption in portfolio selection, there is only one parameter, and it is not difficult to show, as in many places including Steuer, Qi, and Hirschberger (2007), that $E[U(r_p)]$ is a function of the mean and variance of r_p in the form of

$$E(r_p) - \frac{1}{t} V(r_p) \tag{3}$$

where t is a risk tolerance parameter. With (3) concave and increasing, all potentially optimizing solutions of Model (2), with (3) substituted for $E[U(r_p)]$, can be obtained by computing all efficient (E, V) mean–variance combinations that occur in the following two-objective program:

$$\begin{aligned} &\max\ E = E(r_p) \\ &\min\ V = V(r_p) \\ &s.t. \quad \sum_{i=1}^{n} w_i = 1 \\ &w_i \geq 0 \quad \text{for all } i \end{aligned} \tag{4}$$

Recognizing that the two objectives are to be optimized simultaneously, an (E, V) combination is efficient if and only if it is not possible to improve one of the criteria without deteriorating the other. Putting (4) into practice, we have

$$\begin{aligned} &\max\ E = \sum_{i=1}^{n} \mu_i w_i \\ &\min\ V = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \sigma_{ij} w_j \\ &s.t. \quad \sum_{i=1}^{n} w_i = 1 \\ &w_i \geq 0 \quad \text{for all } i \end{aligned} \tag{5}$$

where μ_i is the expected return of the ith security (that is, of the r_i random variable), σ_{ii} is the variance of r_i, and the σ_{ij}, $i \neq j$, are the covariances of the random variables r_i and r_j over the holding period. In bi-criterion format, this is the famous mean–variance model of Markowitz (1952), and as prescribed by Markowitz, the approach is as follows:

1. Specify Model (5) with all of its required μ_i, σ_{ii} and σ_{ij} values.
2. Solve Model (5) for all efficient (E, V) combinations and the fund allocation solution vector $\mathbf{w}$, as a function of V, pertaining to them. Methods for doing this go back to Markowitz (1956).
3. Display the efficient (E, V) combinations in the form of a graph, called the efficient frontier.
4. Have the investor select from the efficient frontier his or her most preferred (E, V) combination.
5. For this (E, V) combination, retrieve from the $\mathbf{w}$ of Step 2 the specific portfolio composition corresponding to the V of the selected (E, V) combination. Provided all has been carried out accurately, this then is the investor's optimal portfolio.

As seen, the efficient frontier is central to the approach. This is because the efficient frontier displays precisely all efficient (E, V) combinations. That is, if a particular fund allocation vector can potentially be an optimal solution of Model (2), its (E, V) combination will be on the efficient frontier, and conversely, if a particular fund allocation vector cannot be an optimal solution of Model (2), its (E, V) combination will not be on the efficient frontier.

The success of Markowitz's mean–variance approach is often attributed to its mathematical tractability, but there are other reasons. One is that the approach allows different investors to have different optimal portfolios. Another is that, because one's optimal portfolio is usually only recognized as such after seeing that everything else is worse, the approach's efficient frontier lets one see the "everything else." However, a caveat comes with the approach.

While the evolution of Model (5) represents considerable achievement with regard to theory, the model is in fact a monster with regard to its demands for data. That is, for an upcoming holding period, the model needs n expected returns, n variances, and $(n^2 - n)/2$ covariances. This is a lot of information, and there may be no good way to get all of it. Hence there is a legitimate worry that errors in the values used for at least some of these quantities will propagate through Model (5) and affect the resulting "optimal" solution.

Fortunately, the σ_{ii} and the σ_{ij} do not create any especial difficulties as they are readily estimated from historical data and tend to be stable from holding period to holding period. However, as brought into sharp relief by Best and Grauer (1991), the μ_i are a different story. Not only are the μ_i lacking in the persistence of the variances and covariances (see DeMiguel & Nogales, 2009; Kan & Smith, 2008; Siegal & Woodgate, 2007), but as shown in Chopra and Ziemba (1993), at a risk tolerance of 50, errors in the μ_i are about 11 more serious than errors of the same relative size in the variances,[1] and about twice that factor with regard to errors of the same relative size in the covariances, with the "times factor" becoming larger as risk tolerance increases. Thus, because of the lessened effect of errors in the variances and covariances, and because errors in the variances and covariance are likely to be of smaller relative sizes than in the expected returns, the implication for research, as pointed out by Chopra and Ziemba (1993), is that first emphasis is to be placed on the expected returns and that is done in this paper.

As a result, we will be comparing two situations. In the first we assume the default case in which the investor has no particular information about the individual security expected returns for the upcoming holding period and is forced to utilize mean returns derived from historical data when setting the μ_i in Step 1 of the Markowitz approach. Following through on the remaining steps of the Markowitz approach, the investor is ultimately guided to a specific fund allocation vector in Step 5. After implementing the specific fund allocation vector only to find out that the actual returns for the holding period differ from the mean returns, the investor then realizes that the portfolio return for the holding period is only $\tilde{E}$ (where $\tilde{E}$ is the dot product of the vector of actual returns and the specific fund allocation vector).

The other situation imagines that the investor is able to come into possession of a forecast for the upcoming period so that the values entered for the μ_i in Step 1 of the Markowitz approach are the returns that will actually occur over the holding period. In this case, guided by the Markowitz approach, the investor will wind up in Step 5 with a fund allocation vector that is optimal for the problem (because of the exact match between the actual security returns and those forecasted). In this case let us say that the portfolio return is $\hat{E}$ (where $\hat{E}$ is the dot product of the vector of actual returns and the optimal fund allocation vector). Thus, the difference between the two portfolio returns (i.e., the difference between $\hat{E}$ and $\tilde{E}$) is the value of information (Copeland & Weston, 1983) which can assist investors in deciding whether more μ_i information beyond the historical should be sought and at what expenditure.

This paper thus proposes a methodology for conducting value of information calculations in portfolio selection. In Section 2 the idea of an efficient frontier is reviewed. In Section 3 a two-security example along with graphs, similar to as in Markowitz (1959, pp. 83–86) and Frances and Archer (1971, pp. 62–70), is provided for initial illustration purposes. Section 4 comments in more detail about the value of information. Section 5 extends the concepts of the paper to n assets. Section 6 illustrates an application drawn from the Taiwan Stock Exchange, and Section 7 provides a discussion of it. Section 8 describes the strategy of parametric quadratic programming that is used to carry out the computations of the paper, and Section 9 closes the paper with concluding remarks.

2. Efficient frontier

Consider n assets whose upcoming holding period returns $r_1, \ldots, r_n$ are described by a probability density function with mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ and covariance matrix $\boldsymbol{\Sigma} = \|\sigma_{ij}\|$.

There are two main ways in which to express the bi-objective formulation of Model (5) in the form of a single-objective program so as to compute the efficient frontier. One is to solve for the value E that maximizes expected portfolio return for a given upper bound on portfolio return variance (i.e., risk) V as in

$$\begin{aligned} & E = \max \sum_{i=1}^{n} \mu_i w_i \\ & s.t. \quad \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \sigma_{ij} w_j \le V \\ & \sum_{i=1}^{n} w_i = 1 \\ & w_i \ge 0 \quad \text{for all } i \end{aligned} \tag{6}$$

Model (6) is a quadratically constrained program due to the $w_i \sigma_{ij} w_j$ terms in the variance constraint. Actually, Model (6) is a very intuitive way of modeling a portfolio problem as many clients have risk levels they do not wish to violate. Various second-order cone solvers (as in Cplex (2013)) can be applied to Model (6), so solutions to the formulation are always within reach this way. By utilizing various V-values to compute their corresponding E-values, the locus of all (E, V) combinations obtained yields the efficient frontier of the problem and the w_i-values of all points along the efficient frontier enable the specification of the fund allocation solution vector $\mathbf{w}$ as a function of V.

The other way of expressing Model (5) in the form of a single-objective program is to find the value of V that minimizes portfolio return variance for a given lower bound on expected portfolio return E as in

$$\begin{aligned} & V = \min \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \sigma_{ij} w_j \\ & s.t. \quad \sum_{i=1}^{n} \mu_i w_i \ge E \\ & \sum_{i=1}^{n} w_i = 1 \\ & w_i \ge 0 \quad \text{for all } i \end{aligned} \tag{7}$$

By utilizing various E-values to compute their corresponding V-values, the locus of all (E, V) combinations obtained also gives the same efficient frontier of the portfolio problem. Since Model (7) is a quadratic program, for which it is generally easier to find solvers, it is often used to find the efficient frontier. In this study, however, Model (6) is used, because it shows the expected portfolio return at each specific level of risk, which is more meaningful in expressing the value of information.

3. An $n = 2$ illustrative example

Assume two risky assets whose historical holding period returns r_1 and r_2 are believed to be described by the bivariate distribution whose mean vector is (8, 6) and whose covariance matrix has elements $\sigma_{11} = 0.75$, $\sigma_{22} = 0.48$, and $\sigma_{12} = 0$. Utilizing this information in Model (6) gives us in the default situation the following formulation:

$$\begin{aligned} & E = \max \; 8w_1 + 6w_2 \\ & s.t. \quad 0.75w_1^2 + 0.48w_2^2 \le V \\ & w_1 + w_2 = 1 \\ & w_1, w_2 \ge 0 \end{aligned} \tag{8}$$

Fig. 1 shows the feasible region in decision space of this problem, which is the intersection of an ellipse (with its interior points) and the straight line of $w_1 + w_2 = 1$ in the first quadrant. For V in Model (8) equal to 0.2927, or 36/123 to be exact, there is only one point C = (0.3902, 0.6098), or (48/123, 75/123) to be precise, in the feasible region. For this V, point C is then the fund allocation

[1] According to Chopra and Ziemba (1993, p.10), most large institutional investors have risk tolerances in the 40–60 range.

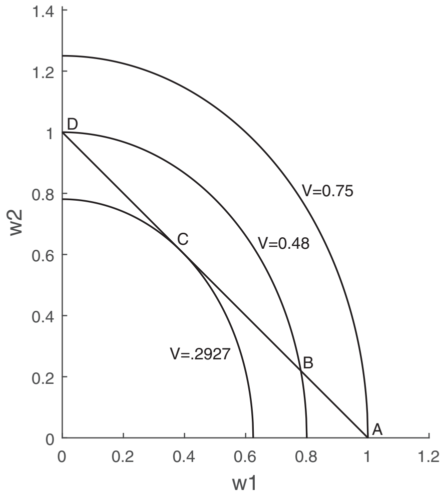

Fig. 1. Graph of the example in decision space.

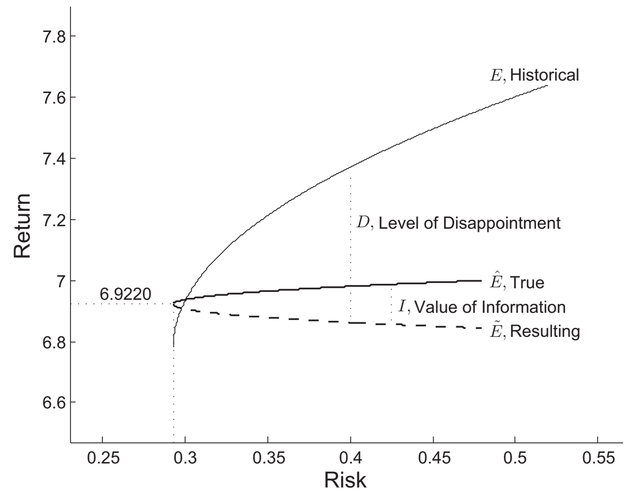

Fig. 3. Values of information and levels of disappointment in the example.

vector that solves (8) whose expected portfolio return $E = 6.7805$, or 834/123 to be exact.

The problem is obviously infeasible when the risk level is less than 0.2927. As risk V increases, the ellipse expands, and the feasible region extends from point C towards points A and D along the line $w_1 + w_2 = 1$. The largest feasible region is the whole line segment AD which occurs when portfolio risk reaches 0.75. Since $\mu_1 > \mu_2$ in this example, the fund allocation vector that solves Model (8) is always at the lower endpoint of the portion of AD that is feasible, regardless of the value of V. A simple mathematical derivation produces, as a function of V, the optimal fund allocation solution vector $\mathbf{w}^*$ and associated expected portfolio return E as follows:

$$\mathbf{w}^* = \frac{(0.96 + \sqrt{4.92V - 1.44},\ 1.5 - \sqrt{4.92V - 1.44})}{2.46}$$
$$E = \frac{8.34 + \sqrt{4.92V - 1.44}}{1.23} \tag{9}$$

for $V \in [0.2927,\ 0.75]$. The efficient frontier of this model, since the individual security expected returns come from historical data, is indicated by the curve labeled "E, Historical" in Fig. 2. Starting from the smallest attainable risk level of 0.2927, expected portfolio return increases from 6.7805, at a decreasing rate, to 8 (not shown) where the risk level is 0.75.

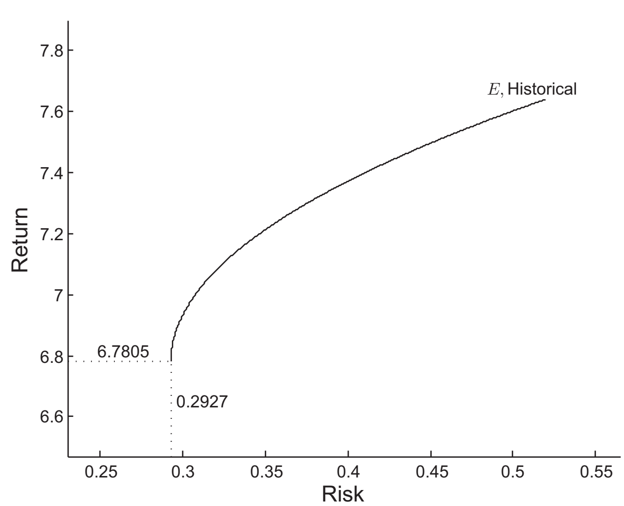

Fig. 2. Efficient frontier of Model (8).

4. Value of information

Now, suppose someone comes to you, claiming that he is a financial guru, and that he can give you an expected return vector, designated $\hat{\boldsymbol{\mu}}$, to be used in place of $\boldsymbol{\mu}$ in Model (6) for the model to yield an optimal fund allocation vector $\hat{\mathbf{w}}^*$ for the upcoming holding period. You are interested in this vector of expected returns in order to determine how much it might be worth to you. Actually, the best $\hat{\boldsymbol{\mu}}$ to be used in this situation is the vector of returns that will actually occur in the upcoming holding period. This would represent the outer limit on forecasting ability as no forecast can be better than that.

If in our example the guru tells you $\hat{\boldsymbol{\mu}} = (8, 6)$, this information has no value at all because this is the vector, from historical data, that you would be using by default if you had no information. However, if the guru tells you $\hat{\boldsymbol{\mu}} = (6.8, 7)$, this information has value because with it one can find his or her optimal fund allocation vector (different from the default case).

By applying Model (8), with $\hat{\boldsymbol{\mu}} = (6.8, 7)$ instead of $\boldsymbol{\mu} = (8, 6)$, one obtains, as a function of V, the optimal fund allocation solution vector $\hat{\mathbf{w}}^*$ and associated portfolio return $\hat{E}$ as follows:

$$\hat{\mathbf{w}}^* = \frac{(0.96 - \sqrt{4.92V - 1.44},\ 1.5 + \sqrt{4.92V - 1.44})}{2.46}$$
$$\hat{E} = \frac{8.514 + 0.1\sqrt{4.92V - 1.44}}{1.23} \tag{10}$$

for $V \in [0.2927,\ 0.48]$. Here, the efficient frontier of Model (8), but with $\hat{\boldsymbol{\mu}}$, is drawn in Fig. 3 and labeled "$\hat{E}$, True". This is because, with $\hat{\boldsymbol{\mu}}$ being the ultimate forecast, "$\hat{E}$, True" is the true efficient frontier, but this curve would not be known without the guru.

Recall from Eq. (9) that $\mathbf{w}^* = ((0.96 + \sqrt{4.92V - 1.44})/2.46,\ (1.5 - \sqrt{4.92V - 1.44})/2.46)$. This is the fund allocation solution vector resulting from $\boldsymbol{\mu} = (8, 6)$. But since $\boldsymbol{\mu} = (8, 6)$ is incorrect, $\mathbf{w}^*$ will not produce as high a portfolio return (for a given V) as $\hat{\mathbf{w}}^*$ does when using the $\hat{\boldsymbol{\mu}} = (6.8, 7)$ given by the guru. Thus, when substituting $\mathbf{w}^*$ into the true objective function "max $6.8w_1 + 7w_2$", we then obtain, as a function of V, the portfolio return of $\tilde{E}$, where

$$\tilde{E} = \frac{8.514 - 0.1\sqrt{4.92V - 1.44}}{1.23} \tag{11}$$

for $V \in [0.2927,\ 0.48]$. Plotting (11) as a function of V produces in Fig. 3 the dashed curve labeled "$\tilde{E}$, Resulting". Without the guru, this curve would not be known either, but it describes reality if the investor uses historical data rather than the guru's $\hat{\boldsymbol{\mu}}$.

Recapping, what we have thus far is this. An investor (could be a mutual fund, pension fund, hedge fund) agrees to take on a level of risk not to exceed V in the upcoming holding period. If the investor utilizes the guru's $\hat{\boldsymbol{\mu}}$ vector, then the investor will implement $\hat{\mathbf{w}}^*$ from Eq. (10) and achieve the portfolio return of the point corresponding to V on "$\hat{E}$, True". If, however, the investor chooses to turn down the guru's $\hat{\boldsymbol{\mu}}$ and go with historical data, then the investor will implement $\mathbf{w}^*$ from Eq. (9), which in general will be incorrect, and only wind up with the portfolio return of the point corresponding to V on "$\tilde{E}$, Resulting", with the difference between these two curves being the value of information.

The difference between the true portfolio return and the resulting portfolio return, $I = \hat{E} - \tilde{E}$, is the loss due to making an incorrect decision, which is also the value of the information provided by the guru. At the smallest risk level of $V = 0.2927$, where there is only one point in the feasible region, one will not make a wrong decision, so the loss due to making a wrong decision there is zero, and the corresponding value of information is also zero. Therefore, if one is paranoid about making a wrong decision, the investor should choose the portfolio of lowest risk, as no mistake can be made there. But as risk increases, the loss due to making a wrong decision, or the value of information, characteristically[2] increases. In the example the value of information is given by $I = \hat{E} - \tilde{E} = 0.2\sqrt{4.92V - 1.44}/1.23$ which yields a minimum value of zero at $V = 0.2927$, a maximum value of 0.1561 at $V = 0.48$, and an overall average over the V-range of "$\hat{E}$, True" of

$$\bar{I} = \frac{1}{0.48 - 0.2927}\int_{0.2927}^{0.48} (0.2\sqrt{4.92V - 1.44}/1.23)dV = 0.1041$$

Being under the illusion in the default case that the investor's expected portfolio return is given by the point corresponding to V on "E, Historical", only to find out in reality that one's portfolio return is given by the point corresponding to V on "$\tilde{E}$, Resulting", the vertical distance between these two points in terms of expected return would be the investor's level of disappointment. That is, when applying the incorrect fund allocation vector $\mathbf{w}^*$, you are expecting a portfolio return of $\boldsymbol{\mu}^T\mathbf{w}^* = E = (8.34 + \sqrt{4.92V - 1.44})/1.23$. But since $\hat{\boldsymbol{\mu}} = (6.8, 7)$, your resulting portfolio return is $\hat{\boldsymbol{\mu}}^T\mathbf{w}^* = \tilde{E} = (8.514 - 0.1\sqrt{4.92V - 1.44})/1.23$. The difference between E and $\tilde{E}$, designated D, shows the level of your disappointment. This value varies with the level of risk. In the example, the level of disappointment is given by

$$D = E - \tilde{E} = (-0.174 + 1.1\sqrt{4.92V - 1.44})/1.23$$

Note that D can be negative, indicating that your portfolio return is more than your erroneous expectation. For example, at $V = 0.2927$, the most risk averse case, the level of disappointment is $D = 6.7805 - 6.9220 = -0.1415$, which is the most negative one. The curves E and $\tilde{E}$ intersect at $V = 0.2978$. If for some reason being disappointed is to be avoided, an investor would then wish to choose a risk level between 0.2927 and 0.2978 in this problem.

5. $n > 2$ assets

The two-asset example can be generalized to n assets. Let us first consider the default case in which the $\boldsymbol{\mu}$-vector of Model (6) is composed of mean returns obtained from historical data. In this model, the fund allocation solution vector is $\mathbf{w}^*$ and this causes the investor to believe that his expected portfolio return is given by $\boldsymbol{\mu}^T\mathbf{w}^*$. While $\mathbf{w}^*$, when $n = 2$, is the straight line in decision space that runs from the point of minimum variance to the point[3] that maximizes $\sum_{i=1}^n \mu_i w_i$, when $n \geq 3$, the line running between the two points is no longer in general a single straight line. It is in general a piecewise linear path, or in other words, a connected path of linear line segments. It is noted that the points along the path at which one segment connects to another are called "turning points" as the path changes direction at these points. How a piecewise linear path comes about is described as follows.

In Model (6), the constraint $\sum_{i=1}^n \sum_{j=1}^n w_i \sigma_{ij} w_j \leq V$ defines an ellipsoid in n-dimensional space. Its intersection with the hyperplane $\sum_{i=1}^n w_i = 1$ is another ellipsoid, but in $(n-1)$-dimensional space. At the smallest attainable risk level, the $(n-1)$-dimensional ellipsoid degenerates to one point, its center. Starting at this point, the ellipsoid expands with risk level V. Since the objective function is linear, the fund allocation solution vector $\mathbf{w}^*$ starts out on a ray emanating from the point of minimum variance as when $n = 2$, but before getting too far will likely hit a $w_i \geq 0$ constraint. As long as the objective function of Model (6) can increase, $\mathbf{w}^*$ will follow the constraint. Then maybe another such constraint is encountered, and $\mathbf{w}^*$ will follow it as long as it is possible to increase the value of the objective function. The process continues in this way, with $\mathbf{w}^*$ in effect zigzagging through the feasible region, until the point that maximizes the objective function of Model (6) is reached.

Now when the guru-supplied objective function coefficient vector $\hat{\boldsymbol{\mu}}$ is utilized, Model (6) becomes Model (12)

$$\begin{aligned}
&\hat{E} = \max \sum_{i=1}^n \hat{\mu}_i w_i \\
&s.t. \quad \sum_{i=1}^n \sum_{j=1}^n w_i \sigma_{ij} w_j \leq V \\
&\sum_{i=1}^n w_i = 1 \\
&w_i \geq 0 \quad \text{for all } i
\end{aligned} \tag{12}$$

where the fund allocation solution vector that solves Model (12) as a function of V, following earlier notation, is denoted $\hat{\mathbf{w}}^*$. But because Model (6) has objective function coefficient vector $\boldsymbol{\mu}$, and Model (12) has objective function coefficient vector $\hat{\boldsymbol{\mu}}$, assuming $\boldsymbol{\mu} \neq \hat{\boldsymbol{\mu}}$, $\hat{\mathbf{w}}^*$ will head away from the point of minimum variance in a different direction and encounter constraints differently than $\mathbf{w}^*$ in search of the point that maximizes over V the value of Model (12)'s objective function $\sum_{i=1}^n \hat{\mu}_i w_i$. Thus, with the two vectors generally becoming further apart as V increases, the value of information, $I = \sum_{i=1}^n \hat{\mu}_i(\hat{w}_i^* - w_i^*)$, typically increases, but of course there can be, as alluded to in footnote [2], intervals of V along which I decreases due to differences between the piecewise linear paths of $\mathbf{w}^*$ and $\hat{\mathbf{w}}^*$.

The level of disappointment is the difference between the expected portfolio return calculated from historical data, $\sum_{i=1}^n \mu_i w_i^*$, and the portfolio return $\sum_{i=1}^n \hat{\mu}_i w_i^*$ experienced by the investor in reality, which can be expressed as $\sum_{i=1}^n (\mu_i - \hat{\mu}_i) w_i^*$. While the $(\mu_i - \hat{\mu}_i)$ are constants as a function of V, the weighted sum $\sum_{i=1}^n (\mu_i - \hat{\mu}_i) w_i^*$ is not, and can easily be increasing along one of the linear segments of $\mathbf{w}^*$ and decreasing along the next. Thus it cannot be predicted what a given level of disappointment curve will do at a given point and depends upon the problem.

6. Taiwan stock market application illustration

For the $n = 2$ illustrative problem of Sections 3 and 4, all can be carried out manually and shown graphically. But when, as discussed in Section 5, a problem involves more than two securities and perhaps takes on additional constraints, the path in decision space that corresponds to the problem's efficient frontier will in almost all certainty be more complicated. When this is the case,

[2] In more complex problems, increases in the value of information may not be strictly monotonic with V as there may be subintervals of V (seen later) in which the value of information decreases slightly before resuming its typically increasing trend.

[3] We are assuming here that the two points of this sentence are each unique.

access to tools like Matlab (2013) and mathematical programming software become necessary .

In this section, to demonstrate the concepts of the paper on a larger problem, a real problem from an emerging market, in which the path is indeed piecewise linear and price volatility can cause the value of information to be high, is taken from the Taiwan Stock Exchange.

Taiwan is a newly industrialized country, and its rapid economic development starting in the 1990s won it the title of one of "The Four Dragons of Asia". Established in 1960, the Taiwan Stock Exchange (TSE) is in charge of the listing and trading of securities. Currently, there are over 800 companies on the TSE. Because data on Taiwan stocks are required for the problem being set up, Thomson Reuters Datastream is used as our data source. But for countries other than the U.S. and Canada, Datastream is only complete from 1988 onward. Selecting only Taiwanese stocks that are complete since 1988, this then allows us to draw annual return data for our experiments from the period January 1, 1989 to December 31, 2014. While there are 44 such stocks, two are dropped (China General Plastics and Namchow Chemical) because of outlier price behavior. If anything, the dropping of the two securities only causes the results of our problem to be less pronounced than if they were included.

The 42 remaining companies, as listed in Table 1, are all in traditional industries, such as textiles, foods, cement, plastics, construction, pulp, and so on. Now, assume that you are the manager of a $2 billion Taiwanese mutual fund and that the 42 stocks of Table 1 constitute your approved list of securities for investment. You of course do not have to invest in them all, but you cannot go outside of the list.

Because how much to set aside for research is always a difficult decision, suppose that, with regard to the 42 securities, you wish to determine what your fund's value of information was last year (for 2014). In other words, you wish to determine how much money you would have left on the table had you used historical means for the expected returns in Model (6) versus a guru-supplied $\hat{\boldsymbol{\mu}}$ as the objective function coefficient vector in Model (12).

To determine this, imagine yourself back at January 1, 2014, and that you would like to use all 25 years of previous historical data for computing the mean returns of the 42 securities and the covariance matrix $\boldsymbol{\Sigma}$ associated with these securities over the period. This gives us the 25-year mean returns and standard deviations (covariances not shown) listed in columns two and three of Table 1.

Solving Model (6) for many values of V produces the Historical efficient frontier labeled as such in Fig. 4(a). Despite the fact that variance is used in Model (6), and also in Model (12), observe that standard deviation is on the horizontal axis. This is normal. While theory and computation are typically conducted in terms of variance, results in portfolio analysis are commonly displayed in terms of standard deviation for greater interpretability. We follow this practice from this point on in this paper.

7. Discussion

From the Historical frontier of Fig. 4(a), the smallest attainable portfolio standard deviation is 0.2961 (or 29.61 percent) at which point expected portfolio return is 0.1206 (or 12.06 percent). As anticipated, the smallest attainable portfolio standard deviation is less than the standard deviation of any of the 42 companies, but at 29.61 percent the effect of portfolio diversification is muted because the average of the correlation coefficients embedded in the covariance matrix of the problem is 0.5702 (as opposed to around 0.20 for the S&P 500).

Table 1
For the 42 stocks, mean return and standard deviation information over the period 1989–2013, return data for 2014 (column five), and sequence numbers to indicate when stocks enter and leave portfolios as we move up the Historical efficient frontier (column four).

	1989–2013		Sequence	2014
	Mean	Stdev	numbers	return
1. Asia Cement	12.0391	0.344983	1, 9	7.4230
2. Cathay Real Estate Dev	12.3569	0.457145		−4.5047
3. Cathay Finan Holdings	12.8283	0.516924		5.5160
4. Chang Hwa Bank	6.2430	0.445567		4.2038
5. Cheng Loong	21.4448	0.516737	13, 16	−5.3715
6. Chia Hsin Cement	6.6512	0.404714		−3.9228
7. China Chemi & Pharm	15.3401	0.581060		−11.0858
8. China Dev Finan Holdgs	20.9496	0.809674		16.8413
9. China Steel	12.2560	0.375606		2.0717
10. China Wire & Cable	13.4782	0.556276		−10.2578
11. Chung Hwa New Century	6.6054	0.464730		−4.9085
12. Far Eastern Depart Stores	17.7044	0.563518	1, 2	−3.2217
13. Far Eastern Finan Holdgs	13.0503	0.435786		4.1346
14. First Financial Holdgs	6.4791	0.448159		10.2657
15. Formosa Chem & Fibre	18.2970	0.416389	2, 14	−17.5843
16. Formosa Plastics	17.2128	0.397592	8, 11	−7.9871
17. Formosa Taffeta	13.5948	0.365130		−10.3191
18. Goldsun Dev & Construct	13.5345	0.546943		−10.8881
19. Grand Pacific Petrochem	15.1054	0.720537		−22.5087
20. Greatwall Enterprises	19.9376	0.637004	7, 17	39.2065
21. Hua Nan Finan Holdings	7.3682	0.460543		9.0621
22. LCY Chemical	16.0920	0.449873		−55.1796
23. Lien Hwa Indl.	11.5796	0.399889		10.6683
24. Nan Ya Plastics	18.9411	0.440844		−2.3871
25. Pacific Construction	18.3111	0.862419		38.9123
26. San Fang Chemical	19.4491	0.613384		2.8483
27. Shihlin Elec & Engine	9.8421	0.425025		7.7174
28. Taiwan Cement	13.5755	0.437849		−1.2233
29. Taiwan Glass Industries	10.5789	0.325998	1, 8	−25.1987
30. Taiwan Pulp & Paper	7.8612	0.503937		−11.1953
31. TSRC	19.9180	0.526613	6, 19	−16.4105
32. Taiwan Tea	18.6398	0.671788		−24.8869
33. Tatung	13.1615	0.748537		8.0797
34. Teco Elec & Machinery	11.4713	0.354611	1, 3	−8.7969
35. USI	13.6723	0.482036		−12.7775
36. United Micro Eltn.	25.0325	0.723298	1,	23.1765
37. Universal Cement	11.6736	0.436670	1, 10	−3.9040
38. Ve Wong	23.0758	0.938055	16, 20	−0.4825
39. Walsin Lihwa	15.1108	0.598652		4.1854
40. Wei Chuan Foods	21.2306	0.655154	5, 18	−51.6898
41. Yuen Foong Yu Paper	8.3663	0.403238		−5.7510
42. Yulon Motor	18.5310	0.583911	1, 13	−12.6645

Note the sequence numbers in the fourth column of Table 1. The securities with no entries in this column never appear in any portfolio along the piecewise linear path of $\mathbf{w}^*$ corresponding to the Historical frontier. Of those that have entries, moving along $\mathbf{w}^*$ in the direction away from the point of minimum standard deviation, the number before the comma means that that security enters the portfolios of the piecewise linear path at the beginning of that segment, and the number after the comma means that that security leaves the piecewise linear path of portfolios at the end of that segment. Consider the sequence numbers of Company 16. The first means that the company enters the portfolios of $\mathbf{w}^*$ at the beginning of the eighth segment and departs from the portfolios of $\mathbf{w}^*$ at the end of the eleventh segment. Note that Company 36 does not have a second sequence number. This is because Company 36 is in all portfolios along the piecewise linear path from start to finish. As a result, Company 36 is the only security left in $\mathbf{w}^*$ when $\mathbf{w}^*$ reaches the maximum expected return point. With the largest sequence number being 20, this means that the piecewise linear path of this particular $\mathbf{w}^*$ has 20 segments.

In the same way, but with the fifth column of Table 1 as the objective function coefficient vector in Model (12), the piecewise linear path of $\hat{\mathbf{w}}^*$ corresponding to the True efficient frontier in

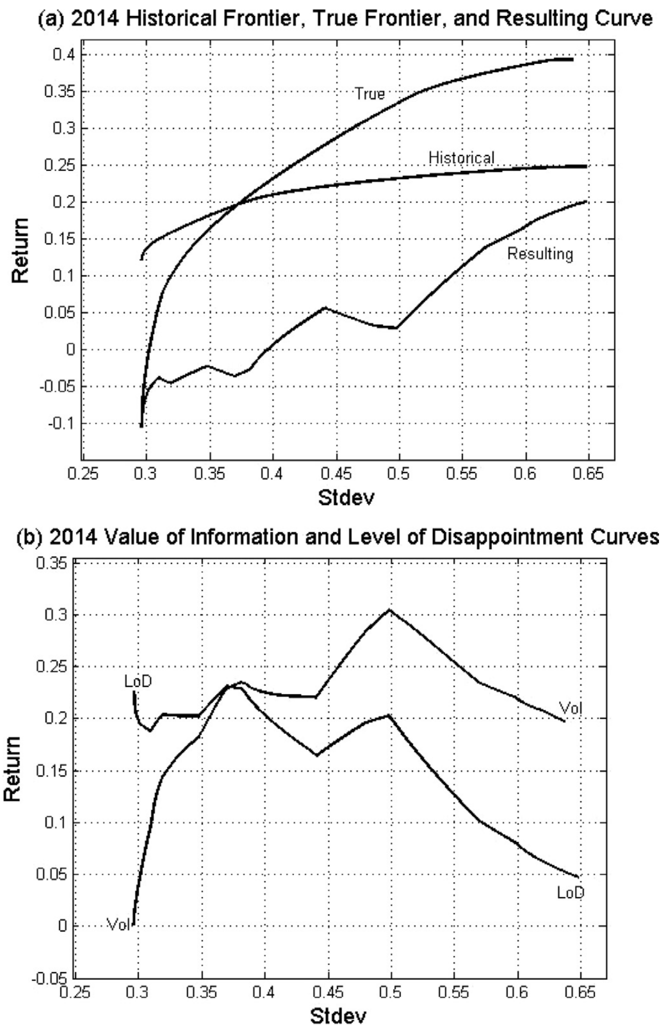

Fig. 4. With reference to the 25 years of immediately previous historical data, (a) the 2014 Historical efficient frontier, True efficient frontier and Resulting curve, and (b) the 2014 Value of Information (VoI) and Level of Disappointment (LoD) curves computed from the frontiers and curve of (a).

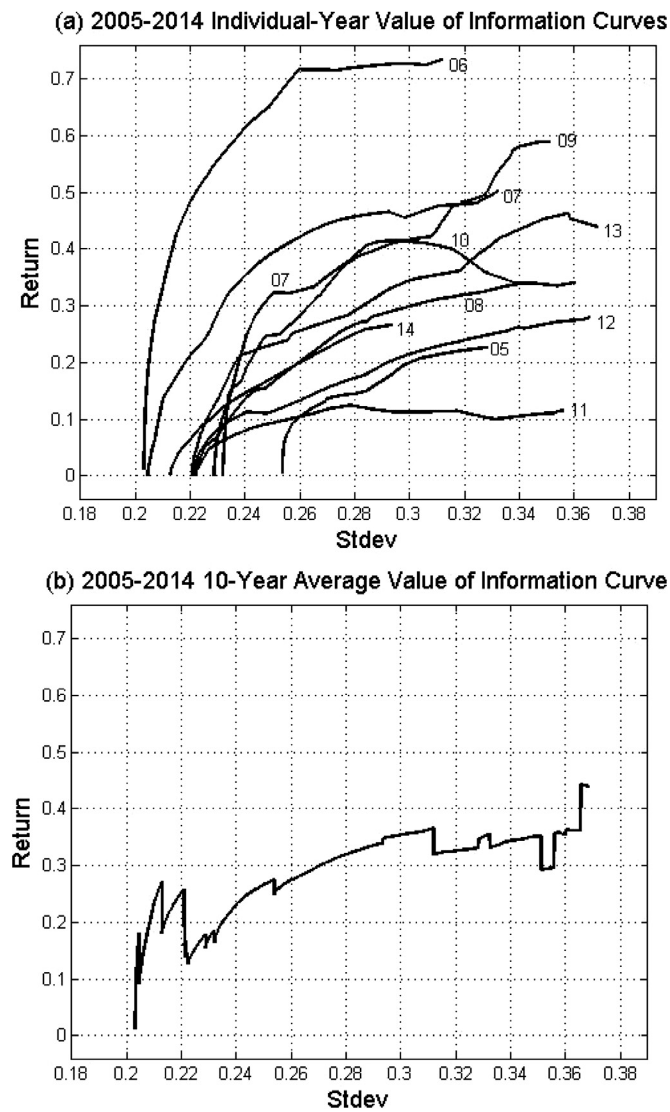

Fig. 5. With reference to 15 years of historical data and an upper bound of 0.10 on all securities in all cases, (a) individual-year Value of Information curves for the years 2005–2014, and (b) average of these curves over the 10-year period.

Fig. 4(a), has 18 segments (sequence numbers not shown). There is no reason why the piecewise linear paths of a historical and a true efficient frontier need involve exactly the same securities or have the same number of segments.

While a given segment of $\mathbf{w}^*$ produces its own smooth portion of a resulting curve, the reason for the bumpiness of the Resulting curve in Fig. 4(a) is that $\mathbf{w}^*$ has 20 segments, meaning that there are 19 turning points all of which can potentially create a kink in the Resulting curve. Because the Resulting curve is to be subtracted from the True frontier and again subtracted from the Historical frontier, the bumpiness of the Resulting curve easily carries over to the Value of Information (VoI) and Level of Disappointment (LoD) curves as seen in Fig. 4(b).

Note that the 2014 Value of Information curve of Fig. 4(b), other than for near the minimum standard deviation, is essentially over 20 percent for most of its run. For a $2 billion fund, this is 400 million dollars plus. We observe that the VoI curve hits its peak at a standard deviation value just shy of 0.5 because the True and Resulting curves have their greatest vertical separation there. The LoD curve has its lowest value at 0.65 because the Historical and Resulting curves are closest together at this standard deviation value. However, before attempting to attribute too much to the curves of any single year, we need to check on at least two things.

One is that our models so far have been unrealistic in the sense that 100 percent is the upper bound on the amount that can be invested in any security. The other is that using 25 years of history may be too much. To address these items, let us take another look at the problem to calculate what the value of information has averaged over the last 10 years using for each year the previous 15 years of historical data and upper bounds of 10 percent on all securities. To accommodate the new upper bounds, the "$w_i \geq 0$" constraints of Models (6) and (12) are changed to "$0 \leq w_i \leq UB$" where UB, in this instance, is 0.10. With these years of historical data and upper bound values, we have the graphs of Fig. 5.

In Fig. 5(a) we have, under the new conditions, the ten individual Value of Information curves for the years 2005–2014. The top curve, for instance, is the Value of Information curve for individual year 2006. Fig. 5(b) portrays an "average" of the 10 curves. But since the individual curves have different starting points and are of different lengths, we need a convention for how an average of such curves is computed. Consider the standard deviation value of 24 percent. From Fig. 5(a) we see that only 9 of the 10 individual Value of Information curves are operational at this value. Thus, we only average the heights of these 9 Value of Information curves to obtain the height of the average Value of Information curve at this value of standard deviation. At, for instance, 28 percent, all 10 are averaged, but at 34 percent, only 6 are averaged.

As seen, the average Value of Information results of this experiment are not much different from the results of the first experiment when only one year was run and all security upper bounds were at 100 percent. Other than for being close to the minimum standard deviation, the results are mostly over 20 percent where the majority of the curves are involved in the average. In fact, in all of the experiments that we have conducted with reasonable values for (i) upper bounds on the securities, (ii) the number of years of historical data to be employed, and (iii) the number of years to be averaged, results have always been consistent with a 20 percent or more figure, already seen twice in this paper, for all values of standard deviation other than when minimal risk is taken. Thus the \$400 million plus figure mentioned earlier for a \$2 billion mutual fund is not off the mark. Thus, it would probably be imprudent for a manager of a large mutual fund to not spend a few million a year on research in an attempt to capture some of the huge amount of money that otherwise gets left on the table.

The 2014 Level of Disappointment curve in Fig. 4(b) has values over 15 percent for most of its length, indicating that in 2014 you would have been disappointed for earning at least 15 percentage points of return less than you would have expected. But after standard deviation of about 0.55, the curve drops of rapidly. However, since almost anything can happen in any given year, one is not to be swayed by any single year Level of Disappointment curve.

In Fig. 6(a) we have the individual-year Level of Disappointment curves, and in Fig. 6(b), following the same averaging procedure as before, we have the 2005–2014 average Level of Disappointment curve. In the middle range of standard deviation, where readings from most years are present, the average Level of Disappointment is around zero, indicating that over the long haul your levels of disappointment tend to average out. This is anticipated. The only reason there are irregularities at the ends of the average curve is that there are not enough observations there. Even though one's average Level of Disappointment may be zero, the individual Level of Disappointment curves show how varied one's disappointments can be year-to-year, positive to possibly large extents in some years and negative to possibly large extents in others.

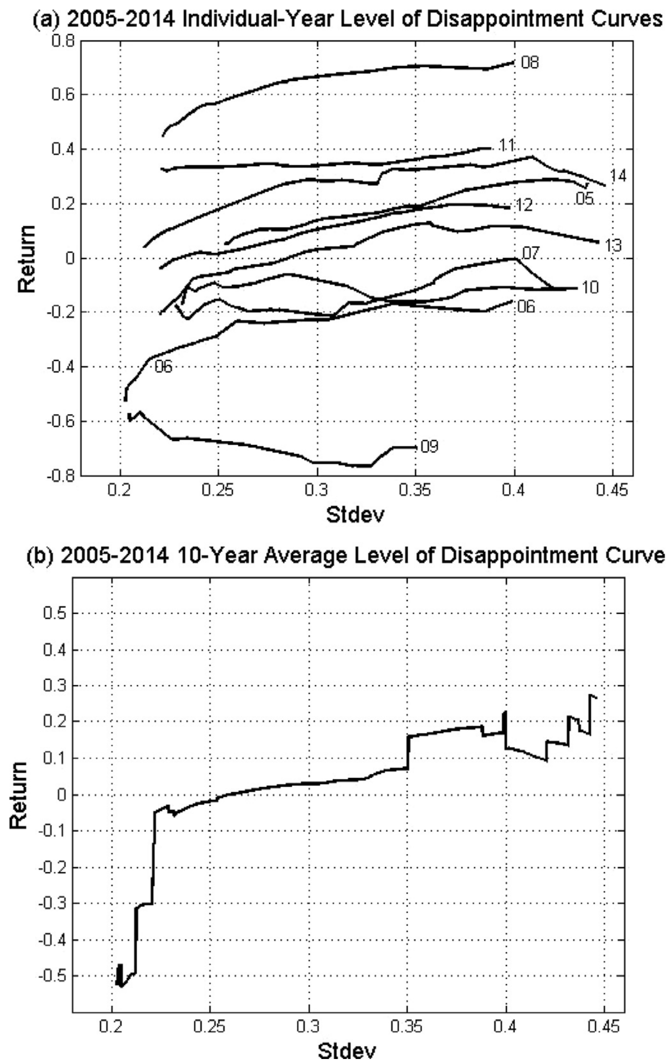

Fig. 6. With reference to 15 years of historical data and an upper bound of 0.10 on all securities, (a) individual-year Level of Disappointment curves for the years 2005–2014, and (b) average of these curves over the 10-year period.

8. Use of parametric quadratic programming

In contrast to the smoothness of the True and Historical efficient frontiers, the Resulting curve, other than in small problems, can be anticipated to exhibit bumpiness. The bumpiness is created by the fact that the fund allocation solution vector $\mathbf{w}^*$ that is optimal for Model (6) is inserted into Model (12) for which it is not optimal. At this point we note that the image of any linear line segment in decision space is a parabolic line segment in *mean–variance* space. That is, each different linear line segment in decision space creates a curved (i.e., parabolic) line segment in *mean–variance* space from a different parabola.[4] Since a parabola in *mean–variance* space is a hyperbola in *mean–standard deviation* space, then each different linear line segment in decision space creates a curved (i.e., hyperbolic) line segment in *mean–standard deviation* space from a different hyperbola. Not for computation as this is all done in terms of variance, but for presentation, this last statement is relevant to us in that all graphs of our application illustration are portrayed with standard deviation on the horizontal axis.

When a piecewise linear path is efficient in a problem, the corresponding hyperbolic line segments in *mean–standard deviation* space will generally blend into one another, sequentially, in a continuously differentiable fashion.[5] Hence the smoothness of the Historical and True efficient frontiers. However, if a piecewise linear path is not efficient in a problem, the hyperbolic line segments will generally not blend smoothly into one another, with each turning point in decision space very well creating a kink in *mean–standard deviation* space. This is the situation when $\mathbf{w}^*$ is placed into Model (12), and this is where the bumpiness of the Resulting curve comes from.

Because the Resulting curve is subtracted from the True and Historical frontiers to obtain the Value of Information and Level of Disappointment curves, respectively, the bumpiness of a Resulting curve is transmitted to these curves as well. Thus, because of the bumpiness phenomenon, it is necessary to solve Models (6) and (12) for many values of V for all curves to be represented accurately as in Figs. 4–6. One particular area where $\mathbf{w}^*$ for many values of V is needed is in the vicinity of where a True frontier and Resulting curve depart from one another. One or both of the curves can be very steep at this point. Steepnesses can occur at other locations, and generally one has no way of knowing where in

[4] As long as no two linear line segments come from the same straight line.

[5] Although kinks are possible, they are rare in problems after more than a few securities.

advance. Consequently, it would be a great relief to be able to simply obtain the solutions of Models (6) and (12) for thousands of values of V, so one would not have to worry about when and where more effort is needed. However, doing this by conventional means, that is, by repetitively calling for each value of V a quadratically constrained solver such as from Cplex from a tool like Matlab, would be quite time consuming, not only because of the cumulative time taken by the solver, but also because of the overhead involved in the calling process.

Instead, we use an algorithm of parametric quadratic programming, in this case the CIOS code from Hirschberger, Qi, and Steuer (2010). This enables us to save considerable time so that the computational requirements of the methodology can be handled in a much more efficient way. An advantage of CIOS is that it can compute in one run a full mathematical specification of the efficient frontier of a problem in very little time (less than 0.04 seconds on average for a problem of 50 securities).[6] By a full mathematical description, we mean a complete description of the piecewise linear path and all parabolic (or hyperbolic) segments. A further advantage of CIOS is that it does not require the covariance matrix to be positive definite which it never is when the number of time periods used to construct the covariance matrix is the same or less than the number of securities being addressed, a common occurrence.

The method of parametric quadratic programming makes use of the fact that the efficient frontier of either Model (6) or (12) is the set of all nondominated solutions of

$$\max \bar{E} = \sum_{i=1}^{n} \bar{\mu}_i w_i$$
$$\min V = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \sigma_{ij} w_j$$
$$s.t. \quad \sum_{i=1}^{n} w_i = 1$$
$$0 \le w_i \le UB \quad \text{for all } i \tag{13}$$

depending upon whether $\bar{\mu}_i$ is μ_i or $\hat{\mu}_i$, and that CIOS is designed to solve (13) for all such nondominated solutions. With the ability of the parametric quadratic programming procedure of CIOS to quickly compute any Historical or True efficient frontier, the task is to ascertain from the output of CIOS, as fast as possible, the expected returns of all $(\bar{E}, V)$ points along the two frontiers that have as their second components the Vs in the list to be processed. Also, it is necessary to compute all of the portfolios along the piecewise linear path $\mathbf{w}^*$ that generates the Historical frontier that have for their Vs the Vs in the same list. All of this is necessary to have points on the Historical, True, and Resulting curves for the same (large) set of Vs to facilitate the vertical subtractions and averagings necessary to produce the graphs of the methodology. Suggested by a method outlined in Qi, Hirschberger, and Steuer (2009), the method here is as follows:

1. Invoke CIOS to compute a mathematical specification of the efficient frontier of interest.
2. From the mathematical specification, note the V-values of the lower and upper endpoints of each contributing parabolic segment.
3. For each in the list of V-values to be processed, do the following:
   3.1 Identify the parabolic segment whose lower and upper V-endpoint values bracket the V-value being processed.
   3.2 Retrieve from the mathematical specification the a_i values that define the identified parabolic segment in the form of
   $V = a_2\bar{E}^2 + a_1\bar{E} + a_0$
   3.3 Rearrange to form
   $0 = a_2\bar{E}^2 + a_1\bar{E} + (a_0 - V)$
   and then apply the quadratic formula to solve for $\bar{E}$.
   3.4 If it is the Historical frontier being processed, let E^{lo} and E^{up} be the expected returns of the lower and upper endpoints of the identified parabolic segment. Then, for the efficient portfolio associated with the V being processed on the piecewise linear path of the Historical frontier, we have
   $$\mathbf{w}_v = \frac{E^{up} - \bar{E}}{E^{up} - E^{lo}}\mathbf{w}^{lo} + \frac{\bar{E} - E^{lo}}{E^{up} - E^{lo}}\mathbf{w}^{up}$$
   where $\mathbf{w}^{lo}$ and $\mathbf{w}^{up}$ are the turning points in decision space pertaining to the lower and upper endpoints of the identified parabolic segment in criterion space, and $\bar{E}$ is the value obtained from the quadratic formula of the previous step.

The above is the procedure for deriving the needed solution information for a given frontier. Considering the study of this paper, there are 10 years and two frontiers (Historical and True) per year for a total of 20 frontiers. With Cplex taking about 40 seconds to solve a given Model (6) or (12) for 1000 values of V, this is 800 seconds in total. Compared to this, CIOS along with a code written in Matlab to implement the above procedure for all 20 frontiers takes only about 8 seconds, which, at about 1 percent of the Cplex alternative, is much more satisfactory from an operational point of view.

9. Conclusions

The problem of portfolio selection can be viewed as the problem of determining the optimal proportions of capital to invest in a set of assets for the purpose of either maximizing expected return for a given upper bound on risk, or minimizing risk for a given lower bound on expected return. However, accurate estimates of individual asset expected returns, the procuring of which generally poses a considerable challenge, are crucial to the process of properly obtaining these proportions. With investors having to commonly proceed with estimates that are not fully accurate, the purpose of the paper is to examine in detail what happens inside of a portfolio selection problem when this is the case. This causes the paper to focus on the concept of the value of information and to introduce the concept of levels of disappointment. Just as Markowitz (1952), Frances and Archer (1971), and others have used small numerical/graphical examples to illuminate their discussions in portfolio selection, we have used an example in the same style to illustrate our work in portfolio selection. In our work with the value of information and levels of disappointment, there are two major findings.

First, the value of information, which is zero at the lowest attainable risk level, generally increases to high levels with risk thereafter. The value of information is of course a key reference point in helping an investor judge how much to spend on research. Second, different from the value of information, the level of disappointment at the lowest attainable risk level is not necessarily zero. It can be positive or even negative, indicating that an expected return higher than anticipated can be experienced.

The whole idea is applied to a case from the Taiwan stock market. An interesting result is that an investor will earn over the long haul what is expected, as indicated by a value around zero for the level of disappointment, but there can be severe differences year to year. What is important and attractive is that an investor

[6] All computations in this paper were conducted on an i7-3770 3.4 gigahertz computer.

could earn on average 20 or more extra percentage points of return per year if sufficient information is provided. Such is quite possibly true about other developing markets, which deserve further research. Possible extensions to the work of this paper that could be interesting would be to apply the methodology to problems with semi-continuous variables (Calvo, Ivorra, & Liern, 2011) and to models with additional objectives as the topic of multiple criteria portfolio management has been attracting increasing attention (Xidonas, Mavrotas, Krintas, Psarras, & Zopounidis, 2012).

References

Abdelaziz, F. B., Aouni, B., & El-Fayedh, R. (2007). Multi-objective stochastic programming for portfolio selection. *European Journal of Operational Research, 177*(3), 1811–1823.

Aouni, B., Colapinto, C., & La Torre, D. (2014). Financial portfolio management through the goal programming model: Current state-of-the-art. *European Journal of Operational Research, 234*(2), 536–545.

Arenas Parra, M., Bilbao Terol, A., & Rodríguez Uría, M. V. (2001). A fuzzy goal programming approach to portfolio selection. *European Journal of Operational Research, 133*(2), 287–297.

Ballestero, E., & Romero, C. (1996). Portfolio selection: A compromise programming solution. *Journal of the Operational Research Society, 47*(11), 1377–1386.

Beraldi, P., Violi, A., & Simone, F. D. (2011). A decision support system for strategic asset allocation. *Decision Support Systems, 51*(3), 549–561.

Best, M. J., & Grauer, R. 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*(2), 315–342.

Bilbao-Terol, A., Pérez-Gladish, B., Arenas-Parra, M., & Rodríguez-Uría, M. V. (2006). Fuzzy compromise programming for portfolio selection. *Applied Mathematics and Computation, 173*(1), 251–264.

Caballero, R., Cerdá, E., Muñoz, M. M., Rey, L., & Stancu-Minasian, I. M. (2001). Efficient solution concepts and their relations in stochastic multiobjective programming. *Journal of Optimization Theory and Applications, 110*(1), 53–74.

Calvo, C., Ivorra, C., & Liern, V. (2011). The geometry of the efficient frontier of the portfolio selection problem. *Journal of Financial Decision Making, 7*(1), 27–36.

Chopra, V. K., & Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, Winter, 6–11.

Copeland, T. E., & Weston, J. F. (1983). *Financial theory and corporate policy*. Reading, MA: Addison-Wesley.

Cplex (2013). IBM ILOG CPLEX optimization studio, Version 12.6.

DeMiguel, V., & Nogales, F. J. (2009). Portfolio selection with robust estimation. *Operations Research, 57*(3), 560–577.

Fang, Y., Lai, K. K., & Wang, S. (2008). Fuzzy portfolio optimization. In *Volume 609 of Lecture notes in economics and mathematical systems*. Berlin: Springer Verlag.

Frances, J. F., & Archer, S. H. (1971). *Portfolio analysis*. Englewood Cliffs, New Jersey: Prentice-Hall.

Hirschberger, M., Qi, Y., & Steuer, R. E. (2010). Large-scale MV efficient frontier computation via a procedure of parametric quadratic programming. *European Journal of Operational Research, 204*(3), 581–588.

Huang, C. F., & Litzenberger, R. H. (1988). *Foundations for financial economics*. Englewood Cliffs, New Jersey: Prentice-Hall.

Kan, R., & Smith, D. R. (2008). The distribution of the sample minimum-variance frontier. *Management Science, 54*(7), 1364–1380.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*(1), 77–91.

Markowitz, H. M. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly, 3*(1–2), 111–133.

Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification in investments*. New York: John Wiley.

Matlab (2013). MathWorks, Inc. Natick, Massachusetts, Version R2013b.

Qi, Y., Hirschberger, M., & Steuer, R. E. (2009). Dotted representations of mean–variance efficient frontiers and their computation. *INFOR, 47*(1), 15–21.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica, 20*(3), 431–449.

Shapiro, A., & Philpott, A. (2007). *A tutorial on stochastic programming*. Atlanta, GA, USA: School of Industrial and Systems Engineering, Georgia Institute of Technology.

Siegal, A. F., & Woodgate, A. (2007). Performance of portfolios optimized with estimation error. *Management Science, 53*(6), 1005–1015.

Steuer, R. E., Qi, Y., & Hirschberger, M. (2007). Suitable-portfolio investors, nondominated frontier sensitivity, and the effect of multiple objectives on standard portfolio selection. *Annals of Operations Research, 152*, 297–317.

Xidonas, P., Mavrotas, G., Krintas, T., Psarras, J., & Zopounidis, C. (2012). *Multicriteria portfolio management*. Berlin: Springer-Verlag.