Michigan Business School

Research Papers

A MINIMAX PORTFOLIO SELECTION RULE WITH LINEAR PROGRAMMING SOLUTION

Working Paper #9612-24

Martin Young
University of Michigan Business School

Research Support
•
School of Business Administration
•
University of Michigan
•
Ann Arbor, Michigan 48109-1234

RESEARCH SUPPORT
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

OCTOBER 1996

A MINIMAX PORTFOLIO SELECTION RULE WITH LINEAR PROGRAMMING SOLUTION

Working Paper #9612-24

Martin Young
University of Michigan Business School

A Minimax Portfolio Selection Rule with Linear Programming Solution

Martin R. Young
University of Michigan School of Business
Department of Statistics and Management Science
Ann Arbor, MI 48109-1234
myoung@umich.edu

Abstract

A new principle for choosing portfolios based on historical returns data is introduced; the optimal portfolio based on this principle is the solution to a simple linear programming problem. This principle uses minimum return rather than variance as a measure of risk. In particular, the portfolio is chosen which minimizes the maximum loss over all past observation periods, for a given level of return. This objective function avoids the logical problems of a quadratic (non-monotone) utility function implied by mean-variance portfolio selection rules. The resulting minimax portfolios are diversified; for normal returns data, the portfolios are nearly equivalent to those chosen by a mean-variance rule. Framing the portfolio selection process as a linear optimization problem also makes it feasible to constrain certain decision variables to be integer, or 0–1, valued; this feature facilitates the use of more complex decision making models, including models with fixed transaction charges, and models with Boolean-type constraints on allocations.

1 Introduction

Sharpe (1971) has remarked that "if the essence of the portfolio analysis problem could be adequately captured in a form suitable for linear programming methods, the prospect for practical application would be greatly enhanced". In this paper a portfolio selection principle is introduced which is such that the optimal portfolio is a solution to a simple linear program. The principle is referred to as "Minimax": the optimal portfolio is defined as that one which would minimize the maximum loss (in dollars) over all past historical periods, subject to a restriction on the minimum acceptable average return across all observed time periods. This principle leads to portfolio selections similar to those obtained by the mean–variance selection rule of Markowitz (1991), in the case when returns have a sample distribution which is approximately multivariate normal.

Since linear programming is becoming a standard feature on personal computer spreadsheet programs, this new minimax rule has the potential to make portfolio optimization a tool accessible to any financial manager. Framing the portfolio selection process as a linear optimization problem also makes it feasible to constrain certain decision variables to be integer, or 0–1, valued; this feature facilitates the use of more complex decision making models. For example, a linear–integer programming model can accommodate fixed transaction charges, a cost commonly encountered by portfolio managers. In addition to the minimax rule's computational convenience, the method may also have logical advantages when the returns are non-normally distributed, and when the investor has a strong form of risk aversion.

1.1 The Minimax Portfolio Selection Rule

Suppose data are observed for N securities, over T time periods. Let

$$\begin{aligned}
y_{jt} &= \text{Return on one dollar invested in security } j \text{ in time period } t \\
\bar{y}_j &= \text{Average Return on security } j \\
&= \frac{1}{T}\sum_{t=1}^{T} y_{jt} \\
w_j &= \text{Portfolio allocation to security } j. \\
y_{pt} &= \text{Return on portfolio in time period } t \\
&= \sum_{j=1}^{N} w_j y_{jt} \\
E_p &= \text{Average Return on portfolio} \\
&= \sum_{j=1}^{N} w_j \bar{y}_j \\
M_p &= \text{Minimum return on portfolio} \\
&= \min_t y_{pt}.
\end{aligned}$$

The "Minimax Portfolio" maximizes the quantity M_p, subject to the restriction that E_p exceeds some minimum level, say G, and that the sum of the portfolio allocations does not exceed some total allocation budget, say W. That is, the minimax portfolio minimizes the maximum loss, where loss is defined as negative gain, or, alternatively, maximizes the minimum gain. [1]

It can be seen that the minimax portfolio is given by the solution to the following linear program:

[1]Thus, it might be more descriptive to refer to the rule as a "maximin" portfolio selection criterion; the term "minimax" will be used throughout, though, due to its greater familiarity.

$$\max_{M_p,\mathbf{w}} M_p \tag{1a}$$

subject to

$$\sum_{j=1}^{N} w_j y_{jt} - M_p \geq 0, \quad t = 1, \ldots, T \tag{1b}$$

$$\sum_{j=1}^{N} w_j \bar{y}_j \geq G \tag{1c}$$

$$\sum_{j=1}^{N} w_j \leq W \tag{1d}$$

$$w_j \geq 0, \quad j = 1, \ldots, N \tag{1e}$$

Equations (1b) guarantee that M_p will be bounded from above by the minimum portfolio return; since this is the only constraint on M_p, and since M_p is being maximized, it will take on the value of the maximum minimum return, or the minimum maximum loss.

An equivalent formulation seeks to maximize expected return, subject to a restriction that the portfolio return exceeds some threshold H in each observation period:

$$\max_{\mathbf{w}} \; E = \sum_{j=1}^{N} w_j \bar{y}_j \tag{2a}$$

subject to

$$\sum_{j=1}^{N} w_j y_{jt} \geq H, \quad t = 1, \ldots, T \tag{2b}$$

$$\sum_{j=1}^{N} w_j \leq W \tag{2c}$$

$$w_j \geq 0, \quad j = 1, \ldots, N \tag{2d}$$

The minimax portfolio is defined as being optimal with respect to the dataset $\{y_{jt}\}$. The $\{y_{jt}\}$ may be a set of historical observations; they may also be values simulated from a probabilistic model for future returns. If one lacks both historical data on past returns, and a predictive probability model for future returns, the minimax method will not be applicable. In such a case, one may need to resort to simpler methods of portfolio selection; e.g., equal weighting across all securities.

2 Relation to Mean–Variance Analysis

Markowitz (1952) presented a quantitative approach to portfolio analysis; his prescription remains the dominant technique in use today. Markowitz proposed choosing the portfolio which minimizes *variance*, subject to a restriction on the mean return. The inputs to this procedure are the means for each security, and the covariances between all securities. In terms of observed data $\{y_{jt}\}$, the Markowitz mean–variance portfolio optimization problem is:

$$\min_{\mathbf{w}} \sum_{j=1}^{N} \sum_{k=1}^{N} w_j w_k s_{jk} \tag{3a}$$

subject to

$$\sum_{j=1}^{N} w_j \bar{y}_j \geq G \tag{3b}$$

$$\sum_{j=1}^{N} w_j \leq W \tag{3c}$$

$$w_j \geq 0, \quad j = 1, \ldots, N, \tag{3d}$$

with $s_{jk} = \frac{1}{T-N} \sum_{t=1}^{T} (y_{jt} - \bar{y}_j)(y_{kt} - \bar{y}_k)$ being the sample covariance between securities j and k. Problem (3) is a quadratic programming problem. In this section, we examine the relationship between the minimax portfolio selection rule, and the mean–variance, or Markowitz, selection rule.

2.1 Normally Distributed Returns

The minimax rule uses minimum return as a measure of portfolio volatility. More specifically, the minimax rule considers the *first order statistic* of the portfolio returns, out of a sample of size T, as an estimate of a portfolio's downside risk. The mean-variance rule, on the other hand, uses variance as the measure of portfolio volatility. In the case of normal data, these two measures of risk – minimum return and variance of returns – will be quite similar.

When the security returns are jointly normal, all portfolios are univariate normal, with some mean μ_p and variance σ_p^2. For a variate X with arbitrary density $f(x)$ and distribution function $F(x)$, the expected value of the first order statistic from a sample of size n is given by

$$e_{1:n} = E\left[X_{1:n}\right] = \int_{-\infty}^{\infty} x f(x) \left(1 - F(x)\right)^{n-1} dx. \tag{4}$$

While (4) is not tractable for arbitrary n in the case when $f(x)$ is $\mathrm{N}\left(\mu, \sigma^2\right)$, David (1981) provides an approximate bound:

$$e_{1:n} \geq \mu - \frac{(n-1)\sigma}{(2n-1)^{\frac{1}{2}}} \tag{5}$$

An alternative approximation sometimes used is

$$\frac{e_{1:n} - \mu}{\sigma} \approx \Phi^{-1}\left(\frac{1}{n+1}\right), \tag{6}$$

where Φ^{-1} is the inverse of the standard Gaussian cumulative distribution function.

Using either approximation (5) or (6), it will be the case that for all portfolios with a given mean return, the portfolio with the smallest variance is also the one with the largest (approximate) expected first order statistic, for any sample size n. Thus the minimax portfolio, the solution to the linear program (1a)–(1e), will be similar to the portfolio which solves the quadratic programming problem (3) minimizing portfolio *variance* given a restriction on the mean.

2.2 Non–Normally Distributed Returns

If the returns of the individual securities are not jointly normal, then a portfolio formed from these securities may be non–normally distributed. A non–normal portfolio distribution cannot in general be characterized by just a mean and a variance. It is known that the mean-variance portfolio selection rule can have counter–intuitive behavior given non–normal probabilities on portfolio returns. For example, given the two portfolios A and B described in Table 1, portfolio B is dominated by portfolio A, in that portfolio A has higher return regardless of the future state of nature. Thus, no investor should want to put 100% of his or her holdings in portfolio B. However, portfolio B is "mean–variance efficient", since it has lower variance than portfolio A. Mean–variance analysis, then, would suggest that investing 100% of ones holdings in portfolio B is reasonable, though logic and utility theory suggest otherwise.

[INSERT TABLE 1 HERE]

The reason mean–variance analysis fails in this setting is that what one truly wishes to achieve in a risk–averse portfolio selection rule is the avoidance of *low* returns, whereas mean–variance selection penalizes variation even among high returns.

The minimax rule is defined in terms of representative data, rather than in terms of a predictive probability distribution function. However, the same situation arises in this setting. Given the data on two securities A and B in Table 2, a mean–variance analysis based on observed sample moments would indicate that a portfolio with 100% of its holdings in security B is efficient, on the grounds that it is less risky than a portfolio with 100% of holdings in security A, even though security B's return is always lower than that of security A. On the other hand, the only minimax-efficient portfolio based on the data in Table 2 will be one with 100% holdings in security A.

[INSERT TABLE 2 HERE]

2.3 The Minimum as a Measure of Volatility

For normal portfolio data, the sample variance will be the most efficient estimator of population variation; i.e., it will attain the Cramèr–Rao lower bound for estimation error (Lehmann 1983). However, the lower order statistic, which is the measure of risk used in the minimax selection rule, may prove, under certain distributions, to be a better estimate of the riskiness of an asset or portfolio than is the sample variance, which is the measure used in mean–variance analysis. Consider, for example, a dataset $\{x_1, \ldots, x_n\}$, with the variates x_i following the log–normal distribution $\mathrm{LN}(\mu, \tau)$ with probability density $f(x) = x^{-1}\left(2\pi\tau^2\right)^{-\frac{1}{2}} \exp\left(-\left(\log x - \mu\right)^2/2\tau^2\right)$. The parameter τ is the scale, or risk, parameter of the distribution; a question is, which summary statistic of the x_i is better for estimating τ: the minimum, $x_{1:n}$, or the sample variance, $v = \sum_{i=1}^{n}(x_i - \bar{x})^2/(n-1)$? Neither statistic is *sufficient* for τ, the sufficient statistics for μ and τ being the sample mean and variance of the *logs* of the x_i (see, e.g, Lehmann 1983 for a definition of sufficiency). However, either statistic can be used, together with the sample mean of the x_i, to form a moment estimator of τ.

The mean and variance of the $\mathrm{LN}(\mu, \tau)$ distribution are given by

$$E[X] = \exp^{\mu + \frac{\tau^2}{2}} \tag{7}$$

$$Var[X] = \exp^{2\mu+\tau^2}\left(\exp^{\tau^2} - 1\right); \tag{8}$$

equating sample moments to population moments leads to the moment estimator:

$$\hat{\tau}_v = \left\{\log(1 + v/\bar{x}^2)\right\}^{1/2} \tag{9}$$

based on the statistics $(\bar{x}, v)$.

Alternatively, since the variable $\log(X)$ will have the normal distribution, one can base an estimator on the first order statistic $x_{1:n}$, using the relation 6:

$$\frac{\log(e_{1:n}) - \mu}{\sigma} \approx \Phi^{-1}\left(\frac{1}{n+1}\right), \tag{10}$$

where, again, $e_{1:n}$ is the expected value of $x_{1:n}$. Equations (7) and (10) together imply

$$\frac{1}{2}\tau^2 - \Phi^{-1}\left(\frac{1}{n+1}\right)\tau + \log\left(\mu/e_{1:n}\right) = 0. \tag{11}$$

A moment estimator $\hat{\tau}_1$ based on the sample statistics $(\bar{x}, x_{1:n})$ can thus be defined as the positive solution of the quadratic equation:

$$\frac{1}{2}\hat{\tau}_1^2 - \Phi^{-1}\left(\frac{1}{n+1}\right)\hat{\tau}_1 + \log\left(\bar{x}/x_{1:n}\right) = 0. \tag{12}$$

A simulation study shows that $\hat{\tau}_1$ is frequently superior to $\hat{\tau}_v$ for estimating the scale parameter of the log–normal distribution. Table 3 presents the mean squared error for estimation of τ based on samples of size n from the LN(μ, τ) distribution, for different values of n and τ. For each combination of parameters, 500 simulation replications were performed; the mean squared error was computed as $\text{MSE}(\hat{\tau}) = \sum_{k=1}^{500}(\hat{\tau}_k - \tau)^2/500$, where $\hat{\tau}_k$ is the estimate ($\hat{\tau}_1$ or $\hat{\tau}_v$) at the kth replication, and τ is the true value of the parameter. In all cases examined, $\hat{\tau}_1$, the estimator based on the first order statistic, significantly outperforms $\hat{\tau}_v$, in terms of having lower mean square estimation error. This result suggests that, for log–normal portfolio data, maximizing the lowest order statistic may lead to more stable portfolios out–of–sample than does minimizing variance. This issue is examined in section 6.2 of this paper.

[INSERT TABLE 3 HERE]

3 Utility Analysis

Numerous authors have shown that a rational decision maker makes his or her decisions by maximizing the expected value of a utility function with respect to some probability distribution; Fishburn 1981 offers a survey of these results. Mean–variance procedures for portfolio selection are known to be consistent with maximization of expected utility if either the predictive distribution of returns is jointly normal and the utility function is concave, or if the utility function is quadratic. The latter assumption is a peculiar one, as the quadratic function is non–monotone, and thus implies that one may prefer less money to more. Counter-examples to the mean-variance procedure occur when neither of the above assumptions are appropriate.

It is of interest to discern whether or not the minimax portfolio selection procedure is consistent with an expected utility framework. In fact, the minimax rule can be motivated by the assumption of a utility function $U(y)$ which implies an extreme form of risk aversion. In particular, assume that the utility function takes the form

$$U(y) = y - \exp\left(-a(y - \tilde{y})\right) \tag{13}$$

for some constants $\tilde{y}$ and $a > 0$.

For the sake of simplicity, assume the random variable Y corresponding to future wealth is discrete, taking values in $\{y_1, y_2, \cdots, y_m\}$, with $y_1 < y_2 < \cdots < y_m$, and with associated probabilities $\{p_1, p_2, \cdots, p_m\}$, $\sum_{k=1}^{m} p_k = 1$. Let $\bar{y} = \sum_{k=1}^{m} p_k y_k$ denote the expected value of Y. The expected utility is given by

$$\begin{aligned}
E\left[U(Y)\right] &= \sum_{k=1}^{m} p_k U(y_k) \\
&= \sum_{k=1}^{m} p_k \left(y_k - \exp\left(-a(y_k - \tilde{y})\right)\right) \\
&= \bar{y} - \sum_{k=1}^{m} p_k \exp\left(-a(y_k - \tilde{y})\right) \\
&= \bar{y} - \exp\left(-a(y_1 - \tilde{y})\right) \sum_{k=1}^{m} p_k \exp\left(-a(y_k - y_1)\right). \qquad (14)
\end{aligned}$$

Let a increase toward infinity; then all components of the summation in (14) will become negligible except the first, since $y_k > y_1$ for $k > 1$, and the expected utility will tend to

$$
\begin{aligned}
E[U(Y)] &\to \bar{y} - p_1 \exp\left(-a(y_1 - \tilde{y})\right) \\
&\to \begin{cases} \bar{y} & \text{if } y_1 > \tilde{y}, \\ -\infty & \text{if } y_1 < \tilde{y}. \end{cases}
\end{aligned}
\tag{15}
$$

Thus, according to (15), given a choice between several random portfolios, a decision maker with utility function given by (13) with a very large will choose the portfolio which has the largest expected return $\bar{y}$, subject to the restriction that the minimum possible return of the portfolio is larger than $\tilde{y}$. This is in correspondence with the minimax prescription given in equations (2). In this latter case, the implied probability distribution for the portfolio with weights $\mathbf{w} = (w_1, \ldots, w_N)$ is the empirical distribution of that portfolio; i.e., the discrete uniform distribution on the T points $\{\sum_{j=1}^{N} w_j y_{jt}, \quad t = 1, \ldots, T\}$.

The minimax portfolio selection principle is compatible, therefore, at least in the limit, with an expected utility maximizing framework. Markowitz (1991) (p.293) considered a related "Maximum Loss" portfolio selection principle, which considers only maximum loss and expected return, and found a contradiction with the axioms of expected utility. The reason for this contradiction was that the maximum loss was considered without regard to its probability. In the minimax rule as defined here, two portfolios with equal observed minimum return M_p are also assumed to have equal probabilities for this event to occur in the future; namely, $\frac{1}{T}$. Markowitz' objection would apply to the minimax principle if there were two portfolios with equal observed mean return, and equal observed minimum, but with one portfolio there were two time periods that had the given minimum return. In this case, the risk averse investor would prefer the portfolio without the tie, but the minimax principle would suggest indifference between the portfolios.

The absolute risk aversion (Pratt 1964) implied by the utility function $U(y)$ is given by

$$A(y) = -\frac{U''(y)}{U'(y)} = \frac{a^2}{\exp\left(a(y-\tilde{y})\right)+a}.$$

As a goes to infinity, $A(y)$ tends to zero for all $y > \tilde{y}$, and tends to infinity for $y \leq \tilde{y}$. Thus there is a discontinuity in the implied attitude toward risk. This might be consistent, if not with the attitude of investors, then at least with the attitude of an agent whose contractual target return is $\tilde{y}$.

It should be noted that while this extremely risk-averse utility function $U(y)$ leads to behavior similar to that mandated by the minimax principle, it is not necessarily the case that the minimax rule is appropriate only for such forms of risk aversion. As was seen in section 2.1, the minimax rule is a good approximation to the mean-variance rule for normal data for a large dataset, and this mean-variance rule is in turn optimal for many typical utility functions given a normal distribution on returns. It is plausible to believe that a minimax rule will be a reasonable approximation to the utility maximizing rule for other combinations of utility function and joint probability distribution; this is seen to be the case, for example, in the simulation study presented below in Section 6.2, with log–normally distributed returns.

4 Relation to other LP Approaches

Because of the obvious advantages of casting portfolio analysis in an LP framework, other authors have sought to make such a connection feasible. Sharpe (1971) uses a piecewise-linear approximation to the quadratic term in the mean-variance quadratic program, enabling the problem to be solved using linear programming methods. Stone (1973) derives a similar approximate solution to the mean-variance optimal portfolio in the case when returns are determined by a single–index model, and also introduces an approximate linear-programming solution to the portfolio selection problem when investors have preference for skewness.

The minimax rule, by comparison, does not require a linearization of a nonlinear objective function, and does not require preprocessing of data, such as estimation of beta coefficients. The minimax procedure, though, does require a *dataset* $\{y_{jt}\}$, whereas the Sharpe (1971) and Stone (1973) procedures are based on a *probability model* of returns. In cases where the first and second moments of the joint distribution are determined subjectively, without the benefit of data, the Sharpe and Stone procedures can be applied directly, whereas the minimax procedure could be applied only by generating a random dataset with the specified moments. In most cases, however, a model would be determined by estimating parameters from a past dataset, so that the minimax procedure will offer the more direct approach to utilizing the data. Also, the minimax rule does not suffer from the potential drawbacks of the quadratic or cubic utility functions assumed in the work of Sharpe and Stone.

Yamakazi and Konno (1991) discovered a solution by using an L_1 measure – the sum of absolute deviations about the mean – as a measure of risk; the optimum portfolio is then the solution to a linear programming problem. The notation "L_1" refers to the fact that the portfolio variation is measured in terms of the first power of the absolute deviations. As with the usual quadratic, or L_2 measure, minimizing the L_1 measure of risk implies a non-monotone utility function; namely, a piecewise linear function. The minimax portfolio uses a different norm, the so-called L_∞-norm (Gonin and Money 1989), as the measure of downside risk. The L_1 and the L_∞ risk aversion measures are qualitatively somewhat different. The piecewise linear L_1 measure represents an indifference to risk over any linear region of the piecewise function, while the minimax, L_∞, rule implies, as was shown via the analysis of the absolute risk aversion function, a strong absolute aversion to downside risk. Both rules can be sensible; the question of which rule is more appropriate will depend on the particular investor's attitude toward risk. One logical consequence of using the L_∞ criterion, which may be seen as an advantage or a disadvantage, is that the minimax-rule will be strongly affected by individual outlying values in the historical data. If, out of T observation periods, a certain asset has $T-1$ large positive returns, and one large negative return, the minimax rule will react more strongly to this single negative return than would an L_1 rule. Because this asset will have a high mean value, the minimax optimal rule may still include this asset, though possibly with less weight than would the L_1-optimal allocation.

5 Linear–Integer Programming

An advantage of framing portfolio selection as a linear programming problem is that linear optimization can accommodate constraints that certain decision variables take on only integer values (see, e.g., Murty 1995, for a treatment of mixed linear–integer programming). This feature allows for the portfolio selection model to treat added complexities. Examples of such extensions to the basic model follow.

5.1 Fixed Transaction Charges

Yamakazi and Konno (1991) noted that a potential advantage of solving portfolio optimization problems as a linear program is that many of the allocations in the optimal solution will tend to be 0, by nature of the corner solution property of linear programs. The reason that this can be advantageous is that there may be some fixed cost associated with buying or holding any amount of a given stock. By incorporating integer–valued variables in the decision model, it is possible to explicitly take into account such fixed charges.

Suppose the transaction cost for purchasing a security j involves a fixed charge C_j, plus a variable cost of V_j per unit purchased. If the fixed cost is substantial, then it may be desirable to reduce the incidence of purchases of small lots of securities. That is, if purchasing *any* amount of stock j involves incurring the fixed cost, then one would expect that an economically optimal allocation would avoid purchases of small odd lots of that stock.

To accommodate this possibility in the portfolio selection model, one can add to the model the 0–1 variables

$$z_j = \begin{cases} 1 & \text{if any shares of stock } j \text{ are purchased} \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

That is, if $w_j > 0$, then $z_j = 1$, and if $w_j = 0$, then $z_j = 0$. Suppose the investment $\mathbf{w}$ is to be held for P periods. Then the expected return on the investment, net of transaction charges, will be

$$E = \sum_{j=1}^{N} \left[P w_j \bar{y}_j - C_j z_j - V_j w_j \right].$$

As in problem (2), the objective E is to be maximized, subject to the constraints that the minimum return exceed some threshold H, and that the total allocation not exceed the original stake W:

$$\max_{\mathbf{w},\mathbf{z}} \sum_{j=1}^{N} \left[P w_j \bar{y}_j - C_j z_j - V_j w_j \right] \tag{17a}$$

subject to

$$\sum_{j=1}^{N} w_j y_{jt} \geq H, \quad t = 1, \ldots, T \tag{17b}$$

$$\sum_{j=1}^{N} (w_j + C_j z_j + V_j) \leq W \tag{17c}$$

$$w_j \geq 0, \quad j = 1, \ldots, N \tag{17d}$$

$$z_j = 0 \text{ or } 1, \quad j = 1, \ldots, N \tag{17e}$$

$$w_j \leq L_j z_j, \quad j = 1, \ldots, N. \tag{17f}$$

The quantities L_j in (17f) are large positive numbers, and are introduced to enforce the condition stated in (16). It can be verified that if $w_j = 0$, then the optimal value of z_j will be 0, while if w_j is positive, z_j will necessarily be set to 1. Since each w_j must be less than W, it is sufficient to set L_j to W.

Thus, problem (17) maximizes the return net of transaction costs, subject to the limitation on risk. See Winston 1994 for a general discussion of fixed charge problems in integer programming. The analysis can be extended fairly directly to accommodate piecewise linear transaction charge schedules.

5.2 Logical Side Constraints

By allowing integer–value restrictions on decision variables, it is possible to include logical constraints on the allocation. Examples of logical constraints are given below:

1. "The portfolio may not include both asset j *and* asset k."

2. "The portfolio must hold a_j or more dollars of asset j, *or* a_k or more dollars of asset k."

In general, if one wishes to enforce the logical constraint that *either* $f(\mathrm{w}) \leq 0$ *or* $g(\mathrm{w}) \leq 0$, then one can add the constraints:

$$f(\mathrm{w}) \leq Lz \tag{18a}$$

$$g(\mathrm{w}) \leq L(1-z) \tag{18b}$$

$$z = 0 \text{ or } 1, \tag{18c}$$

(Winston 1994), where L is a suitably large number (larger, that is, than the maximum possible values of $f(\mathrm{w})$ and $g(\mathrm{w})$). Applying this general rule, along with the laws of Boolean algebra, one can devise linear–integer constraints for "exclusive or", "if–then", "if and only if", and other, more complex types of logical restrictions on allocations.

The minimax portfolio selection method, then, can be modified to handle a wide range of modeling complexities. Algorithms have recently been developed for solving quadratic integer programs (see, e.g., Bienstock (1995) and Fletcher and Leyffer (1994)); thus, it may also be possible to accommodate integer constraints in the mean–variance optimization framework. However, linear–integer programming codes can handle much larger problems than can be handled by the quadratic–integer programming methods;

also, software for linear–integer programming is currently more widely available than is software for nonlinear–integer programming.

6 Applications

In this section, the minimax rule is applied to real and simulated datasets, in order to examine the performance characteristics of the portfolio selection method. A small study using real historical data is performed, which provides a visual depiction of the optimization being performed. The simulation is then presented to provide insight concerning the long–run performance of the minimax selection approach.

6.1 International Stock Index Data

A set of historical data on the returns from international stock indices was examined using the minimax and mean–variance portfolio selection rules, to illustrate the difference in the performance characteristics of the two principles. The data consist of monthly returns on the stock indices from Canada, France, Italy, Japan, United Kingdom, United States, and West Germany, from January 1991 until December 1995. The returns data y_{jt} were derived from the stock index prices, P_{jt}, as listed in the Citibase database.[2] The returns y_{jt} were defined as $y_{jt} = (P_{j,t+1} - P_{jt})/P_{jt}$.

The data from the first 30 months – January 1991 to June 1993 – were used to obtain the minimax portfolio and the mean–variance portfolio; the performance of the resulting weights was then examined out–of–sample during the following 30 months of July 1993 to December 1995. The minimax portfolio was obtained by solving problem (1), and the mean–variance portfolio was obtained by solving optimization problem (3). In both cases, the target monthly expected return G was set to 1.0%.

The performance of the portfolios, both in–sample and out–of–sample, is displayed in Figure 1. As expected, the minimax portfolio has a higher "floor" during the estimation

[2]The prices P_{jt} are listed as Citibase variables FPS6CA, FPS6FR, FPS6IT, FPS6JP, FPS6UK, FPS6US, and FPS6WG.

period of January 1991 to June 1993, while the mean–variance portfolio has less variance during this period. The summary statistics for the returns of the two portfolios are given in Table 4. In this example, the minimax portfolio's performance appears to be competitive with that of the mean–variance portfolio.

[INSERT TABLE 4 HERE]

[INSERT FIGURE 1 HERE]

6.2 A Simulation Study

A simulation study was performed to examine the performance of the minimax and mean–variance portfolio rules over a large universe of data realizations. In the simulation, random realizations of a stock market were generated from a joint probability distribution based on a simple index model:

$$y_{jt} = \exp\left(\alpha_j + \beta_j m_t + \sigma_j \epsilon_{jt}\right), \quad j = 1, \ldots, N, \quad t = 1, \ldots, 2T,$$

with the index m_t normally distributed as $\mathrm{N}(\mu_m, \sigma_m)$, and the ϵ_{jt} distributed independently as $\mathrm{N}(0, 1)$. Thus, the marginal distribution of y_{jt} was log–normal for each security j and time t. The common dependence on an index m_t introduced correlations among the N securities. The parameters α_j for the different securities were distributed uniformly over the range $(\underline{\alpha}, \overline{\alpha})$; similarly, the β_j and σ_j were distributed over the ranges $(\underline{\beta}, \overline{\beta})$ and $(\underline{\sigma}, \overline{\sigma})$, respectively.

For each simulation replication, $2T$ periods of data were generated. The observations in the first T periods were used to estimate optimal minimax and mean–variance portfolios, by solving linear program (1) and quadratic program (3) respectively, with an identical constraint value G. The estimated portfolio weights were then evaluated on the second T simulated periods.

The simulations were performed for a number of different parameter values, in order to observe the effects of sample size and joint probability distribution of returns on the relative performances of the two optimization methods. In particular, the number of securities N, number of observations T, the index volatility σ_m, and the range of α values $(\underline{\alpha}, \overline{\alpha})$, were varied across the different simulation blocks. For each combination of parameter values, $(\underline{\beta}, \overline{\beta})$ was fixed at (0.5, 1.5), and $(\underline{\sigma}, \overline{\sigma})$ was fixed at (0.8, 2.0). 500 simulation replications were performed for each combination of parameter values.

Table 5 lists the values for the parameters used in each simulation block, along with the simulation output: the average and standard deviation of the out–of–sample portfolio mean and variance. The performances of the minimax and mean–variance approaches were compared by paired–t tests on the mean and variance. For example, the first two rows of the table indicate that, for the simulation replications performed with $N = 12$, $T = 18$, $(\mu_m, \sigma_m) = (0, 1)$, and $(\underline{\alpha}, \overline{\alpha}) = (-1, 1)$, the minimax portfolios had an average mean of 18.37 while the mean–variance portfolios had an average mean of 18.18; also, the minimax portfolios had an average variance of 37.63, while the mean–variance portfolios had an average variance of 34.74. As can be seen in Table 5, the minimax and mean–variance portfolios had performances that were not significantly distinguishable; p–values for testing for equality of mean and equality of variance between the two methods varied from 0.104 to 0.999, and were generally above 0.25. 500 simulation replications were performed for each set of parameter values, so this result does not appear to be due to lack of statistical power.

[INSERT TABLE 5 HERE]

This finding that the minimax and the mean–variance portfolio selection methods perform similarly on these simulated log–normal data may seem to be in conflict with the simulation results described in Section 2.3, in which the sample lowest order statistic was seen to be a significantly better predictor of future volatility than was the sample variance, for log–normally distributed data. The conflict may be at least partly explained by the fact that both portfolio optimization methods are concerned with minimizing the volatility of the *portfolio*, which is a linear combination of the individual assets. Central limit theorem results suggest that the portfolio will have a distribution which is close to normal, though the individual assets have non–normal distributions. For a portfolio with a finite number of log–normally distributed assets, the distribution will appear somewhere between log–normal and normal in character. For example, Table 6 displays a skewness coefficient for a hypothetical portfolio P generated by $P = \frac{1}{N}\sum_{j=1}^{N} Y_j$ where $Y_j = \exp\left(am + \sqrt{1-a^2}\epsilon_j\right)$, $j = 1, \ldots, N$, $m \sim \mathrm{N}(0,1)$, $\epsilon_j \sim \mathrm{N}(0,1)$.

The marginal distribution of each Y_j is $\mathrm{LN}(0,1)$, and the correlation between $\log(Y_j)$ and $\log(Y_k)$ is $\rho = a^2$ for all $j \neq k$. The skewness coefficient displayed is $(P_{.75} - P_{.50})/(P_{.50} - P_{.25})$, where $P_{.25}$, $P_{.50}$, and $P_{.75}$ are the 25th, 50th, and 75th percentiles respectively for the distribution of P; for symmetric distributions, this coefficient will equal 1.0. The percentiles were computed by Monte Carlo simulation with 5000 replications. The table shows that as N, the number of securities in the composition of P, increases, the skewness decreases, indicating a shift towards normality. Log–normality in the distribution of the portfolios might favor the minimax approach, while mean–variance analysis is optimal for normal data; a distribution between log–normality and normality could lead to comparable behavior for the minimax and the mean–variance optimizers, as was in fact observed.

[INSERT TABLE 6 HERE]

7 Conclusion

Given historical or simulated future returns data on a collection of assets, an optimal "minimax portfolio" can be constructed using linear programming techniques. Under weak conditions, the minimax principle corresponds approximately to an expected utility maximizing principle, with the implied utility function representing an extreme form of risk-aversion.

There are certainly settings in which previously proposed portfolio selection methods will be preferable to the minimax method. For example, if the assets under analysis follow a single index model whose parameters are known to a reasonable degree of accuracy, and if the utility function implied by mean–variance analysis is a reasonable approximation to the investor's true utility function, then the simple portfolio selection rule suggested in Elton, Gruber, and Padberg (1976) will be preferable to the minimax rule, in terms of both ease of computation, and long run performance.

On the other hand, if an investor's utility function is more risk–averse than is implied by mean–variance analysis, or if returns data are skewed, or if the portfolio optimization problem involves a large number of decision variables, including integer valued variables, the minimax rule may provide a sensible approach to portfolio selection.

Table 1: Returns with non–normal predictive distribution

State of Nature	Prob.	Returns	
		Port. A	Port. B
1	.25	0.20	0.10
2	.50	0.40	0.20
3	.25	0.60	0.30

Table 2: Returns with non–normal empirical distribution

Time	Returns	
	Port. A	Port B
1	0.10	0.01
2	0.20	0.02
3	0.10	0.01
4	0.20	0.02
⋮	⋮	⋮
500	0.20	0.02

Table 3: Monte Carlo evaluation of moment estimators for the scale parameter of the log–normal distribution. $\hat{\tau}_v$ is the estimator based on the sample variance, and $\hat{\tau}_1$ is the estimator based on the sample minimum.

μ	τ	n	MSE($\hat{\tau}_v$)	MSE($\hat{\tau}_1$)	P–val.$\star$
0.0	1.5	25	0.1685	0.0697	$p < 0.0001$
0.0	1.5	100	0.0783	0.0306	$p < 0.0001$
0.0	1.5	500	0.0336	0.0178	$p < 0.0001$
0.0	2.0	25	0.5160	0.0968	$p < 0.0001$
0.0	2.0	100	0.2778	0.0488	$p < 0.0001$
0.0	2.0	500	0.1295	0.0304	$p < 0.0001$

$\star$: P–value for paired–t test of hypothesis MSE($\hat{\tau}_1$) = MSE($\hat{\tau}_v$)

Table 4: Performance of minimax and mean–variance optimal portfolios, in–sample and out–of–sample.

In–sample:				
January, 1991 - June, 1993				
	MEAN	VAR	MIN	MAX
Mean–Variance	1.0E-2	4.84E-04	-1.96E-2	9.18E-2
Minimax	1.0E-2	5.76E-04	-1.40E-2	1.03E-1
Out–of–sample:				
July, 1993 - December, 1995				
	MEAN	VAR	MIN	MAX
Mean–Variance	9.2E-3	3.57E-04	-2.75E-2	3.51E-2
Minimax	1.03E-2	2.69E-04	-2.66E-2	3.21E-2

Table 5: Performance of minimax and mean–variance portfolio optimizers on 500 replications of simulated data.

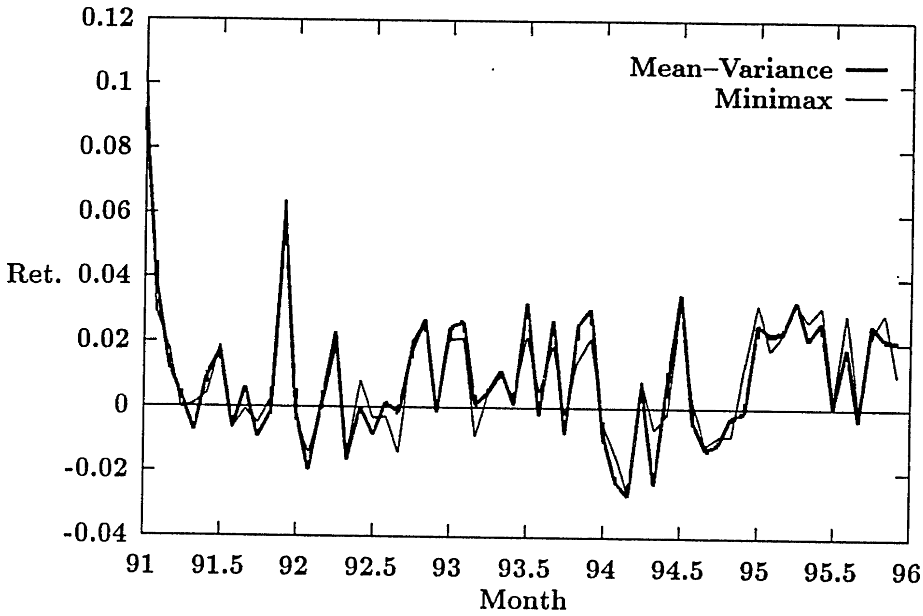
					Minimax		Mean–Variance		P–val.⋆	P–v
N	T	(μ_m, σ_m)	$(\underline{\alpha}, \overline{\alpha})$		Mean	Var.	Mean	Var.	Mean	Va
12	18	(0,1)	(-1.0,1.0)	Avg.	18.37	37.63	18.18	34.74	0.826	0.3
				S.D.	(24.75)	(86.26)	(21.37)	(72.84)		
6	18	(0,1)	(-1.0,1.0)	Avg.	20.13	44.20	21.62	47.61	0.421	0.5
				S.D.	(33.31)	(110.00)	(36.86)	(116.14)		
12	36	(0,1)	(-1.0,1.0)	Avg.	23.76	64.27	21.33	51.66	0.108	0.1
				S.D.	(28.52)	(130.89)	(35.88)	(196.31)		
12	18	(0,1)	(-2.0,1.0)	Avg.	4.43	8.62	4.51	8.37	0.748	0.7
				S.D.	(5.19)	(14.24)	(4.32)	(13.19)		
12	18	(-3,2)	(-1.0,1.0)	Avg.	1.61	4.98	1.60	4.71	0.879	0.5
				S.D.	(4.49)	(16.43)	(4.76)	(17.04)		
16	18	(0,1)	(-1.0,1.0)	Avg.	6.57	12.92	6.57	12.09	0.999	0.5
				S.D.	(9.39)	(30.90)	(5.60)	(17.24)		
10	12	(0,1)	(-1.0,1.0)	Avg.	23.47	46.85	20.30	36.74	0.287	0.2
				S.D.	(73.63)	(215.67)	(38.45)	(108.83)		

⋆: P–value for paired–t test.

Table 6: Skewness of averages of N correlated log–normal variables with cross–correlation ρ.

	N			
ρ	1	5	10	50
0.00	1.96	1.36	1.41	1.14
0.25	1.96	1.45	1.36	1.21
0.75	1.96	1.48	1.33	1.11

Figure 1: Monthly returns from minimax and mean–variance optimal portfolios of international stock indices.



References

Bienstock, D. (1995). Computational study of a family of mixed-integer quadratic programming problems. In *Integer programming and combinatorial optimization*, Berlin, pp. 80–94. Springer.

David, H. (1981). *Order Statistics* (2nd ed.). New York, NY: John Wiley and Sons.

Elton, E., M. Gruber, and M. Padberg (1976). Simple Criteria for Optimal Portfolio Selection. *Journal of Finance 31*, 1341–1357.

Fishburn, P. C. (1981). Subjective expected utility: a review of normative theories. *Theory and Decision 13*, 139–199.

Fletcher, R. and S. Leyffer (1994). Solving mixed integer nonlinear programs by outer approximation. *Mathematical Programming 66*, 327–349.

Gonin, R. and J. Money (1989). *Non-linear L_p-Norm Estimation.* New York, NY: John Wiley and Sons.

Lehmann, E. (1983). *Theory of Point Estimation.* New York, NY: John Wiley and Sons.

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

Markowitz, H. (1991). *Portfolio Selection* (2nd ed.). Cambridge, MA: Basil Blackwell Ltd.

Murty, K. (1995). *Operations Research: Deterministic Optimization Models.* Englewood Cliffs, NJ: Prentice–Hall, Inc.

Pratt, J. (1964). Risk Aversion in the Small and in the Large. *Econometrica 32*, 122–136.

Sharpe, W. (1971). A Linear Programming Approximation for the General Portfolio Analysis Problem. *Journal of Financial and Quantitative Analysis 6*, 1263–1275.

Stone, B. (1973). A Linear Programming Formulation of the General Portfolio Selection Model. *Journal of Financial and Quantitative Analysis 8*, 621–636.

Winston, W. (1994). *Operations Research: Applications and Algorithms* (3rd ed.). Belmont, CA: Duxbury Press.

Yamakazi, H. and H. Konno (1991). Mean Absolute Deviation Portfolio Optimization Model and its Application to Tokyo Stock Market. *Management Science 37*, 519–531.