Theory and Methodology

Using investment portfolio return to combine forecasts:
A multiobjective approach

Mark T. Leung a,*, Hazem Daouk b,1, An-Sing Chen c,2

a Division of Management and Marketing, College of Business, The University of Texas, 6900 North Loop 1604 West, San Antonio, TX 78249-0634, USA
b Department of Finance, Kelley School of Business, Indiana University, Bloomington, IN 47405, USA
c Department of Finance, National Chung Cheng University, Ming-Hsiung, Chia-Yi 621, Taiwan, ROC

Received 1 March 1999; accepted 25 August 2000

Abstract

This study investigates the usefulness and efficacy of a multiobjective decision method for financial trading guided by a set of seemingly diverse analysts’ forecasts. The paper proposes a goal programming (GP) approach which combines various forecasts based on the performance of their previous investment returns. In our experiment, several series of financial analysts’ forecasts are generated by different forecasting techniques. Investment returns on each series of forecasts are measured and then evaluated by three performance criteria, namely, mean, variance, and skewness. Subsequently, these distributional properties of the returns are used to construct a GP model. Results of the GP model provide a set of weights to compose an investment portfolio using various forecasts. To examine its practicality, the approach is tested on several major stock market indices. The performance of the proposed GP approach is compared with those of individual forecasting techniques and a number of forecast combination models suggested by previous studies. This comparison is conducted with respect to different levels of investor preference over return, variance, and skewness. Statistical significance of the results are accessed by bootstrap re-sampling. Empirical results indicate that, for all examined investor preference functions and market indices, the GP approach is significantly better than all other models tested in this study. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Investment analysis; Goal programming; Combining forecasts; Multiobjective decision analysis; Trading strategies

1. Introduction

One of the many contributions of operations research to financial planning is the development of integrated financing and investment models, i.e., the exploitation of investment opportunities in a systematic fashion. In recent years, there is a
growing trend of using multiobjective techniques in this avenue of research. The primary advantage of using multiobjective technique in decision making is, as stated in Spronk (1981), “that most of these (single objective) models and methods are unsuitable for decision situations in which multiple and possibly conflicting objectives play a role, because they are concentrated on the optimal fulfillment of only one objective”. Given this notion, we attempt to explore the possibility of taking the multiobjective approach to solve a typical problem encountered by many financial and investment managers, namely, making investment trading decisions based on a set of potentially incompatibile forecasts supplied by different analysts.

It is surprising that the Nobel prize winning insight of Markowitz (1952) has not penetrated the literature on consensus forecasts. Markowitz’s theory has revolutionized the way people think about portfolios of assets. Trade-offs between competing objectives is in the heart of modern finance and a large portion of operations research. With the increasing expectation of their performance from the companies and investors, financial managers are under the pressure of maximizing the return on the investment while keeping the risk to a reasonable level. In the literature, most of the studies attempted to construct some portfolio which addresses the tradeoff between two concerns, i.e., maximizing the expected return while reducing the variance of the portfolio. In some recent studies, the concept of mean–variance tradeoff was extended to include the skewness of return in multiobjective portfolio selection. The ambition of this paper is to apply the concept of mean–variance–skewness tradeoff in the area of consensus forecasts. Our results show that the benefits gained from this major insight are in parallel to those observed in portfolio management. This paper is unique in that it views analysts forecasts as assets with different characteristics. Combining those “assets” to form an optimal consensus forecast is based on the powerful framework of mean–variance–skewness efficiency. This framework explicitly takes into account utility preferences.

Techniques of forming consensus estimates from a set of forecasts generated by various sources (e.g., financial analysts) are not new and have been well accepted in the academic research community since the seminal work by Bates and Granger (1969). A great number of ways to combine the forecasts have been proposed and tested. Nevertheless, most of these consensus models are based on the concept of forming a linear combination of various forecasts with respect to a particular objective. Traditionally, multiobjective approaches are largely ignored. In addition, majority of the models examined in the literature fails to take into account of the distribution properties of the investment returns. Therefore, in this paper, we propose and investigate a multiobjective decision model which analyzes the properties, such as the mean and variance, of the return distribution. To provide better economic insight, our study also investigates the performance and efficacy of investment trading schemes guided by the proposed multiobjective model.

Unlike the one proposed in this study, most of the portfolio selection models found in the literature involve a set of different financial instruments. Here, we develop an alternate approach for portfolio construction using a single instrument while the issues of expected return, variance, and skewness are still accounted for. It is interesting to point out that this single-instrument portfolio is able to benefit from the variance-reduction effect due to diversification, assuming the current distribution of investment return is a good representative of the one in the future. To better illustrate the logic and functionality of the proposed multiobjective approach, the model is tested on several internationally traded broad market indices.

In summary, the primary focuses of this study are to propose a new approach to combine forecasts and to provide empirical evidence of the performance of our proposed multiobjective approach relative to some previously suggested single-criterion methods. The experiment also examines the magnitude of multiobjective approach’s enhancement on investor preference relative to individual forecasting models (i.e., no forecast combination). The remaining portion of this paper is organized as follows. A literature review of the related concepts and economic justifications are given in Section 2.
Also, a description of our empirical data is included. In Section 3, a summary of the forecasting techniques employed in our experiment which represent the models used by various financial analysts is outlined, followed by a discussion of forecasting results. Then, in Section 4, we present the multiobjective approach to combine forecasts based on their distributions of previous returns and the mathematical programming formulation. Section 5 shows the results of the proposed investment approach and describes a trading simulation used to compare the performance of the multiobjective portfolio with those trading schemes guided by individual forecasting techniques and analyzes the results of the experiment. Section 7 concludes the study.

2. Background

2.1. Combining forecasts

The idea of ‘forecast combination’ was first studied by Bates and Granger (1969). In their seminal work, the authors proposed a linear combination of two forecasts with weights selected to minimize the predicted forecast error variance. This research study was later extended by Newbold and Granger (1974) which formed consensus from more than two forecasts. In a separate study, Granger and Ramanathan (1984) utilized a regression approach to determine the best weights based on a Bayesian methodology. The superior performance of combining forecasts over individual approaches was illustrated in the extensive empirical evaluation conducted by Makridakis and Winkler (1983) and Russel and Adam (1987). Since that period, we have witnessed a dramatic growth in research, both theoretical and applied, validating and extending this seemingly naive approach in forecasting. A comprehensive review of former works in this research stream can be found in Clemen (1989).

Although the literature contains a great diversity of methods to combine forecasts, most studies generate consensus forecasts with respect to a single criterion (e.g., minimum average error or error variance). Reeves and Lawrence (1982) is one of the very few researches that explicitly considered multiple criteria. Essentially, the study proposed and tested a GP model which minimizes total forecast error, positive forecast error over all periods, total forecast error over the recent periods, and maximum forecast error simultaneously. However, a practitioner may find this model inadequate as it does not capture any attribute or preference other than forecast errors. Also, a more general view of the distributional properties of the return on investment are often neglected. In our paper, we look at the practical issue of how to incorporate different analysts’ forecasts into an investment trading model, a problem frequently involved in the basic operations of many financial service institutions. The traditional single objectives of either minimizing forecast error (or error variance) or maximizing return may not suffice here because a financial manager always has to take into account certain risk factors and is willing to accept a smaller return for lower risk levels. Therefore, our multiobjective approach explicitly encompasses the average return, expected variance, and skewness in combining forecasts.

2.2. Portfolio selection with skewness of return

Since Markowitz (1952) proposed a theoretical framework for selecting efficient portfolios, there has been numerous studies examining the issue of diversification of risk. Essentially, the risk level (i.e., variance of the return) is reduced at the expense of expected return. This is usually achieved by forming a portfolio with non-perfectly correlated instruments. In later studies, the importance of the role of skewness in finance was demonstrated by Arditti and Levy (1975) in the pricing of stocks and by Friedman and Savage (1948) in purchasing a lottery ticket. As shown in Arditti (1967), the investor’s preference for more skewness to less is consistent with the notion of decreasing absolute risk aversion. It is because a positive skewness in the return refers to a right-handed elongated tail for the density function. Positive skewness is desirable, since increasing skewness will decrease the probability of large negative returns, while increasing the probability of large positive returns.
The fact that investors care about the skewness of assets returns has been shown by many studies that documents the existence of a risk premium for skewness. For this reason, an optimal portfolio derived by omitting the skewness of each period in a multiperiod model could be an inefficient portfolio. This fact is shown in the numerical example in Lai (1991). He gave a GP procedure that performs portfolio selection based on competing and conflicting objectives—maximizing both expected return and skewness while minimizing the risk associated with the return (i.e., variance). Chunhachinda et al. (1997) applied Lai’s methodology in determining the optimal portfolio which consists of 14 international stock indices. Similar to Lai’s procedure, their GP approach incorporates investor’s preference for skewness. Their results showed that the incorporation of skewness into portfolio selection causes a major change in the construction of the optimal portfolio. The study also provided evidence that investors trade expected return of the portfolio for skewness.

2.3. Data

In the light of the previous literature, it is hypothesized that various measures of the macroeconomic environment may be used as input variables in the construction of prediction models to forecast the stock market index. Table 1 outlines an array of such macroeconomic state variables which are applied to the paper. For the sake of brevity, interested readers can find in the Appendix the description of the economic intuition concerning why these state variables are chosen in this study. The financial and macroeconomic data set used in this study is obtained from TSM data base and Citibase compiled by DSC Data Services, Inc. and Citicorp Economic Database, respectively. The entire data set covers the period from January 1967 to December 1995, a total of 348 months of observations. The data set is divided into two periods: the first period runs from January 1967 to December 1990 (288 months of observations) while the second period is from January 1991 to December 1995 (60 months of observations.) The first period, which is assigned to in-sample estimation, is used for determining the specifications of the models and parameters for the forecasting techniques. It also serves the purpose of validating the estimated models. The second period is reserved for out-of-sample evaluation and comparison of performances between various forecasting models.

To test the versatility and robustness of the multiobjective approach for combining forecasts,

<table>
<thead>
<tr>
<th>Input variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST - Short-term interest rate</td>
<td>First difference of three-month T-bill rate for the US</td>
</tr>
<tr>
<td></td>
<td>First difference of call money rate for the UK and Japan</td>
</tr>
<tr>
<td>LT - Long-term interest rate</td>
<td>First difference of long-term government bond rate for the US</td>
</tr>
<tr>
<td></td>
<td>First difference of twenty-year government bond rate for the UK</td>
</tr>
<tr>
<td></td>
<td>First difference of long-term government bond rate for Japan</td>
</tr>
<tr>
<td>TS - Term structure proxy</td>
<td>First difference of long-term government bond rate minus first difference of three-month T-bill rate for the US</td>
</tr>
<tr>
<td></td>
<td>First difference of twenty-year government bond rate minus first difference of call money rate for the UK</td>
</tr>
<tr>
<td></td>
<td>First difference of long-term government bond rate minus first difference of call money rate for Japan</td>
</tr>
<tr>
<td>R - Lagged index returns</td>
<td>Lagged terms of the continuously compounded excess returns of the broad market index for the different countries, respectively</td>
</tr>
<tr>
<td>CPI - Consumer price level</td>
<td>First difference of consumer price index for the three countries, respectively</td>
</tr>
<tr>
<td>IP Industrial production level</td>
<td>First difference of industrial production for the three countries, respectively</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>R - Returns on index</td>
<td>Continuously compounded excess returns of the broad market index. The excess return for a particular time period is defined as the continuously compounded return minus the risk free rate for the corresponding time period. The broad market indices are chosen as follows: S&amp;P 500 for the US, FTSE 100 for the UK, and Nikkei 225 for Japan</td>
</tr>
</tbody>
</table>
three globally traded broad market indices – S&P 500 for the US, FTSE 100 for the UK, and Nikkei 225 for Japan, are examined in our empirical experiment. The forecast variables are the continuously compounded one month excess returns of these indices. As it is shown in Eq. (1), the excess return on a certain index is defined as the continuously compounded return on the price index minus the riskfree interest rate.

\[ R_t = \ln \left( \frac{P_t}{P_{t-1}} \right) - r_{t-1}, \]  

where \( P_t \) is the price of the stock index traded at period \( t \) and \( r_t \) is the US riskfree (one month T-bill) interest rate in period \( t \). Dividends are ignored for this study. The reason to forecast the excess returns (rather than the index levels or the total returns) is that they provide a measure of how well our models perform relative to the minimum returns gained from depositing the money in a riskfree account. In addition, our study assumes that we are in the position of an American investor. This leads to our adoption of US riskfree T-bill rate, as opposed to the interest rates of other riskfree instruments issued by foreign governments, in deriving the excess returns. This is because an American investor always has a choice of not investing but simply putting the money in his US account which pays the short-term riskfree interest rate.

The independent variables for predicting the index returns are all observable on or before the last day of the month preceding the month to be forecast. For instance, for the prediction of the index return for March 1990, all independent variables must be observable on or before the last day of February 1990. Constructing the data in this manner ensures that the estimation of out-of-sample forecasts will be similar to the practice in the real world. It is because only observable, but not future, data are used as inputs to the forecasting models. The list of potential financial and macroeconomic input variables and the forecast output variables used in our experiment is outlined in Table 1.

### 3. Generating analysts’ forecasts

Before any multiobjective investment portfolio can be constructed, we need to obtain the forecasts from different sources (or analysts). To accomplish this task, we mimic the estimates made by different analysts by creating the forecast series using various forecasting techniques. In a study exploring the effects of combining forecasts generated by different techniques, Makridakis and Winkler (1983) tested 1001 time series and its subset of 111 series with 14 individual forecasting techniques. They found that the variance of overall mean absolute percentage error (MAPE) decreases exponentially with the number of techniques. According to their empirical results, there is a drastic improvement in this performance measure when the number of techniques used in forecast combination increases from 2 to 4. Moreover, the results of the studies conducted by Newbold and Granger (1974), Pindyck and Rubinfeld (1976), and Mahmoud (1984) indicated that combining Box-Jenkins, regression, and exponential smoothing techniques generally produces better forecasts than single model forecasts. Guided by these results, four forecasting techniques – exponential smoothing (AES), autoregressive integrated moving-average (ARIMA), multivariate transfer

---

3 The term “financial analysts” is just an alias to represent the sources of different series of forecasts. In many real-world situations, a decision maker does not have to generate the forecasts using different forecasting techniques. They just simply obtain the different series of forecasts from various analysts. For example, a certain investment banker and brokerage firm has four chief analysts in their North American Division tracking the movements and making recommendations for internet stocks. Each of these chief analysts may have his/her own models which lead to different forecasts (price targets) and ratings. The performance (with respect to the average, variance, and skewness of returns) of these analysts varies over time and there is no analyst who can consistently outperform all others in all criteria. Therefore, it may be more advantageous to combine their forecasts than to solely rely on a single analyst’s estimates. In this case, the actual forecasting technique(s) and model(s) used by an analyst are not major concerns to the decision maker (who combines the forecasts from various analysts). The investment firm as a whole is more interested in using the historical performance of each of these analysts to determine the weights to combine the forecasts.
function (MTF), and artificial neural network (NN) – are chosen to generate series of forecasts as if they are estimated by different analysts. In the following sections, we will briefly summarize the logic of each forecasting technique and the model used in the experiment.

3.1. Exponential smoothing (AES)

Makridakis et al. (1983) and Mabert (1978) described an extension to traditional exponential smoothing model, generally known as adaptive exponential smoothing. This approach continuously evaluates the performance in the previous period and updates the smoothing coefficient. The form of the adaptive exponential smoothing model is similar to that of the simple single exponential smoothing model:

\[
F_{t+1} = \alpha x_t + (1 - \alpha)F_t,
\]

where \(F_t\) is the forecast for period \(t\) and \(x_t\) is the actual observation made in period \(t\), and

\[
\alpha_{t+1} = \frac{E_t}{M_t},
\]

\[
M_t = \beta|e_t| + (1 - \beta)M_{t-1},
\]

\[
E_t = \beta e_t + (1 - \beta)E_{t-1},
\]

\[
e_t = x_t - F_t,
\]

\(\alpha\) and \(\beta\) are parameters between 0 and 1 and \(|\cdot|\) denotes absolute values.

Based on the results of experiment, the values of \(\beta\) are set to be 0.75, 0.90, and 0.95 for the S&P 500, FTSE, and Nikkei prediction models, respectively. These values are determined by accessing the performance of the models in the first 228 months within the in-sample period (from January 1967 to December 1985). The estimated parameters are then validated using the remaining 60 months (from January 1986 to December 1990) within the in-sample period. Since we assume that historical performance provides meaningful information to predict the future, these values for \(\beta\) are used in the forecasting in the reserved out-of-sample period.

3.2. ARIMA

Since the seminal work of Box and Jenkins, ARIMA models have been widely used in the forecasting of economic and financial time series. The general ARIMA\((p, n, q)\) model can be written as

\[
y_t = a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-p} + e_t + b_1 e_{t-1} + \cdots + b_q e_{t-q},
\]

where \(y_t\) is the series of interest and \(e_t\) is the residual. \(n\) indicates the degree of differencing used to obtain stationarity in the model. In our experiment, we follow Box and Jenkins’ three-stage method. 4 aimed at selecting an appropriate model for estimating and forecasting a univariate time series. The procedure yields the following forecasting models for the selected stock indices:

For S&P 500:

ARIMA(1,0,1) or ARMA(1,1)

\[
R_t = -0.18 R_{t-1} + e_t + 0.47 e_{t-1}.
\]

For FTSE:

ARIMA(0,0,1) or MA(1)

\[
R_t = e_t + 0.42 e_{t-1}.
\]

4 Box and Jenkins’ three-stage method aimed at selecting an appropriate model for estimating and forecasting a univariate time series:

1. Identification stage: We use the SARIMA procedure in RATS statistical software to determine plausible models. The SARIMA procedure uses standard diagnostics such as autocorrelation function (ACF), partial autocorrelation function (PACF), and plots of series.

2. Estimation stage: Each of the tentative models is fit and the various coefficients are examined. In this stage, the estimated models are compared using standard criteria such as AIC, SBC, and significance of coefficients.

3. Diagnostic checking stage: SARIMA procedure is used to check if the residuals from the different models are white noise. The procedure uses diagnostics tests such as ACF, PACF, Ljung–Box Q-statistic for serial correlation, and Jarque–Bera normality test.
For Nikkei:
ARIMA(0,0,3) or MA(3)
\[ R_t = \varepsilon_t + 0.39\varepsilon_{t-1} + 0.05\varepsilon_{t-2} + 0.13\varepsilon_{t-3}. \] (10)

Similar to their adaptive exponential smoothing counterparts, the models presented above are estimated based on the first 228 observations within in-sample period (from January 1967 to December 1985). After these estimated models are validated by the last 60 months of observations in the in-sample period, they are used to estimate the out-of-sample forecasts from January 1991 to December 1995.

3.3. Multivariate transfer function (MTF)

A multivariate transfer function model is essentially an ARIMA with added exogenous variables. The general model can be stated as
\[ y_t = a_0 + a_1y_{t-1} + \ldots + a_py_{t-p} + e_t + b_1\varepsilon_{t-1} + \ldots + b_q\varepsilon_{t-q} \]
\[ + \left( \omega_{10} + \omega_{11}L + \ldots + \omega_{1m}L^m \right) x_{1,t-b} + \ldots + \left( \omega_{20} + \omega_{21}L + \ldots + \omega_{2n}L^n \right) x_{2,t-b} + \ldots, \] (11)

where \( x_{1,t-b}, x_{2,t-b}, \ldots \) are exogenous variables lagged for \( b \) periods, \( L \) is the lag operator, \( y \) the dependent variable, and \( \varepsilon \) the residual. The addition of exogenous variables can improve forecasting if these variables help explaining the stock excess returns. Interested readers should refer to Makridakis et al. (1983) for a detailed explanation of the technique. The methodology used in the estimation by multivariate transfer function is the same as that for ARIMA (see Section 3.2). Potential exogenous variables for consideration are listed in Table 1. The procedure (see Footnote 4) yields the following forecasting models for the various stock indices:

For S&P500 of the US:
ARMA(1,1) with short-term interest rates (ST)
\[ R_t = -0.48R_{t-1} + \varepsilon_t + 0.66\varepsilon_{t-1} \]
\[ + \frac{-0.02}{(1 - 0.31L + 0.17L^2 - 0.64L^3)} ST_{t-1}. \] (12)

For FTSE of the UK:
MA(1) with long-term interest rates (LT)
\[ R_t = \varepsilon_t + 0.48\varepsilon_{t-1} \]
\[ + \frac{-0.00 - 0.01L}{(1 + 0.83L - 0.37L^2 - 0.84L^3)} LT_{t-1}. \] (13)

For Nikkei of Japan:
MA(3) with long-term interest rates (LT)
\[ R_t = \varepsilon_t + 0.38\varepsilon_{t-1} + 0.06\varepsilon_{t-2} + 0.12\varepsilon_{t-3} \]
\[ + \frac{-0.01}{(1 - 0.40L + 0.41L^2 - 1.02L^3)} LT_{t-1}. \] (14)

In all cases, Industrial Production (IP) and Consumer Price Index (CPI) did not add predictive value to the model. Like the ARIMA models, the models presented above are estimated and validated using the data from the in-sample period. 60 monthly forecasts for the reserved out-of-sample period are generated using these estimated models.

3.4. Artificial NN

A feedforward layered NN with a single hidden layer is used to perform the forecasting of the index returns.\(^5\) For the operational details of a feedforward network with backpropagation training paradigm, readers should refer to Wasserman (1993).

An imperative issue in designing a network is determining the appropriate number of units in the hidden layer. Unfortunately, there is no definite answer to this question. Based on the results from

\(^5\) Our choice of this construct is justified by Hornik et al. (1989) which proved that a one-hidden-layer feedforward NN is capable of approximating uniformly any continuous multivariate function to any desired degree of accuracy. Hassoun (1995, p. 48), further suggested that “any failure of a function mapping by a multilayer network must arise from inadequate choice of parameters or an insufficient number of hidden nodes”.

our study, we select the network architecture which leads to consistent and reasonable performance in the validation sample period from January 1986 to December 1990. The resulting construct of the network contains twelve input units, ten hidden units, and a univariate output unit for the US and UK models. For the Japan model, the network contains nine input units, six hidden units, and a univariate output unit. The model specifications are as follows:

For S&P 500:
\[
R_t = F(R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, \text{ST}_{t-1}, \text{LT}_{t-1}).
\]  
(15)

For FTSE of the UK:
\[
R_t = G(R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, \text{ST}_{t-1}, \text{LT}_{t-1}).
\]  
(16)

For Nikkei of Japan:
\[
R_t = H(R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, \text{LT}_{t-1}),
\]  
(17)

where \(F(\cdot), G(\cdot), \) and \(H(\cdot)\) are the arbitrary functions deduced by the NNs. The trained networks are then applied to the forecasting of the index returns in the out-of-sample period (January 1991 to December 1995).

4. Theory and models

4.1. Multiobjective investment return model

In earlier sections, we explain the rationale to maximizing the skewness of return for the enhancement of the traditional Markowitz mean–variance portfolio. Thus, our proposed multiobjective approach directly takes into account the mean, variance, and skewness in forming the consensus. Essentially, we adapt and extend Lai (1991) and Chunhachinda et al. (1997) GP formulation, which was originally applied to portfolio selection, to our problem setting. Let \(\tilde{\mathbf{R}} = (\tilde{R}_1, \tilde{R}_2, \tilde{R}_3, \tilde{R}_4)\) denote the distributions of excess return on an index based on the forecasts generated by the four forecasting techniques described in Section 3. The mean of the return with respect to \(\mathbf{R}\) is \(\bar{\mathbf{R}} = (\bar{R}_1, \bar{R}_2, \bar{R}_3, \bar{R}_4)\). Also, let \(\mathbf{W}^T = (w_1, w_2, w_3, w_4)\) be the transpose of the weight vector used to combine the forecasts and \(\mathbf{V}\) be the variance–covariance matrix for \(\mathbf{R}\). Mathematically, we compute the first three moments (mean, variance, and skewness) of \(\tilde{\mathbf{R}}\) as follows:

\[
\text{Mean} = \mathbf{W}^T \bar{\mathbf{R}} = \sum_{i=1}^{4} w_i \bar{R}_i, \quad \text{where}
\]

\[
\bar{R}_i = \frac{1}{N} \sum_{t=1}^{N} R_{it},
\]  
(18)

\[
\text{Variance} = \mathbf{W}^T \mathbf{V} \mathbf{W} = \sum_{i=1}^{4} w_i^2 \sigma_i^2 + \sum_{i=1}^{4} \sum_{j=1}^{4} w_i w_j \sigma_{ij} \quad \text{for } i \neq j,
\]  
(19)

\[
\text{Skewness} = E[\mathbf{W}^T (\tilde{\mathbf{R}} - \bar{\mathbf{R}})]^3 = \sum_{i=1}^{4} w_i^3 s_i^3
\]

\[
+ 3 \sum_{i=1}^{4} \left( \sum_{j=1}^{4} w_i^2 w_j s_{ij} + \sum_{j=1}^{4} w_i w_j^2 s_{ji} \right)
\]

\[
\quad \text{for } i \neq j,
\]  
(20)

where \(\sigma_i^2\) and \(\sigma_{ij}\) are the variance and covariance of the return based on forecasting technique \(i\) whereas \(s_i^3\), \(s_{ij}\), and \(s_{ji}\) are the skewness and coskewness of the return based on forecasting technique \(i\), respectively.

To form the consensus estimates, we use a GP model to combine the forecasts such that the expected return and skewness of return are maximized while the variance (risk level) of the return is minimized. This notion can be represented by the following goal program (GP1):

Maximize \(Z_1 = \mathbf{W}^T \bar{\mathbf{R}}\),

Maximize \(Z_2 = \mathbf{W}^T \mathbf{V} \mathbf{W}\),

Maximize \(Z_3 = E[\mathbf{W}^T (\tilde{\mathbf{R}} - \bar{\mathbf{R}})]^3\),

subject to
\[
\mathbf{W}^T \mathbf{I} = 1,
\]  
(21d)

\[
\mathbf{W} \succeq 0.
\]  
(21e)

A simple but effective way to solve GP1 is to consolidate the various objectives into a single
objective function. Hence, we need to define a compact objective function which captures the values of $Z_1$, $Z_2$, and $Z_3$. Let $d_1$, $d_2$, and $d_3$ be the goal variables which account for the deviations of expected return, variance, and skewness from the aspired levels, $Z_1^*$, $Z_2^*$, and $Z_3^*$, respectively. The aspired level indicates the best case scenario for a particular objective without considering other objectives. Hence, the aspired levels $Z_1^*$, $Z_2^*$, and $Z_3^*$ can be determined by solving three (independent) subproblems relaxed from GP1.

Subproblem GP2A:

Maximize $Z_1^* = W^T \bar{R}$

subject to

$W^T I = 1, \quad W \succeq 0. \quad (22)$

Subproblem GP2B:

Minimize $Z_2^* = W^T V W$

subject to

$W^T I = 1, \quad W \succeq 0. \quad (23)$

Subproblem GP2C:

Maximize $Z_3^* = E[W^T(\bar{R} - \bar{R})]^3$

subject to

$W^T I = 1, \quad W \succeq 0. \quad (24)$

It is interesting to note that the solution of subproblem GP2A is trivial because the goal program will always assign all weights to the highest yielding forecasting technique when other constraints are ignored. On the other hand, solving subproblems GP2B and GP2C do provide meaningful solutions which lead to tighter bounds and the basis for normalization in the objective function.

The objective of the problem can thus be defined as the minimization of deviations from the ideal scenario (which may not be feasible when all objectives are considered simultaneously) set by the aspired levels. Spronk (1981) suggested to use the general Minkowski distance, a metric often used in finance and economics, for the specification of the objective function in goal programming (GP). The computational form of the Minkowski distance is

$$Z = \left\{ \sum_{k=1}^{m} \left( \frac{d_k}{Z_k^*} \right)^p \right\}^{1/p}, \quad (25)$$

where $Z_k^*$ is the basis for normalizing the $k$th goal variable. The value of $Z_k^*$ is set to the corresponding aspired level $Z_k^*$ in this study. In order to allow the investor exhibiting asymmetric preferences toward the mean, variance, and skewness of return, we modify Eq. (25) and introduce parameters $p_1$, $p_2$, and $p_3$ to indicate the relative preference for expected return and skewness, respectively. Also, since $p_i \geq 0$ (for $i = 1, 2, 3$), monotonicity allows us to drop the $p$-root calculation to make the equation more compact. Eq. (26a) shows the adapted version of the Minkowski equation. 6

Now, we can re-formulate GP1 to include the goal variables and our modified objective function. Given the investor’s relative preferences over the mean, variance, and skewness of return, the GP model (GP3) can be stated as:

Minimize $Z = \left\{ \left| \frac{d_1}{Z_1^*} \right|^{p_1} + \left| \frac{d_2}{Z_2^*} \right|^{p_2} + \left| \frac{d_3}{Z_3^*} \right|^{p_3} \right\}$, \quad (26a)

subject to

$W^T \bar{R} + d_1 = Z_1^*$ \quad (26b)

$W^T V W - d_2 = Z_2^*$ \quad (26c)

$E[W^T(\bar{R} - \bar{R})]^3 + d_3 = Z_3^*$ \quad (26d)

$W^T I = 1 \quad (26e)$

$d_1, d_2, d_3, \quad W \succeq 0. \quad (26f)$

6 It should be pointed out that, unlike the GP objective functions used in the studies by Lai (1991) and Chunhachinda et al. (1997), the additive form of Eq. (25) does not create any bias because the values of $(d_k/Z_k^*)$ for $k = 1, 2, 3$ are of comparable scale. If the goal variable $d_k$ is not normalized with respect to the basis $Z_k^*$, the optimization will put excessive weight on those objectives with larger magnitudes, resulting in a biased solution. For the same reason, the preference parameters $p_1$, $p_2$, and $p_3$ do not need to be rescaled.
Since the solutions of the relaxed subproblems are at least as good as the solution attained by GP3, where all objectives \( Z_1, Z_2, \) and \( Z_3 \) are considered simultaneously, the values of the goal variables \( d_1, d_2, \) and \( d_3 \) are always non-negative. In other words, the goal variables represent the amount of underachievement with respect to the best scenario. The optimal weight set \( W = w_1, w_2, w_3, w_4 \) thus forms the basis for constructing a multiobjective portfolio.

4.2. Portfolio-based trading approach

The traditional trading approach which is guided by a single forecasting technique often leads to a unique trading pattern, and requires an investor to choose a specific forecasting model over the others. However, there is no guarantee that the performance of a certain model is consistent from period to period. Eliminating the need of an investor to follow the forecasts set by a single forecasting technique, our proposed portfolio-based trading approach divides the investment into four different components. Each component is independently guided by one of the four forecasting techniques (NN, AES, ARIMA, MTF) and has its own trading pattern. Constructing a portfolio in this manner can possibly alleviate the impact resulted from an unexpected performance deterioration of a single forecasting model.

On the other hand, previous studies, such as Makridakis and Winkler (1983) and Russell and Adam (1987), propose and test a number of models to combine multiple forecasts. However, the forecasts are combined in a way that produces a single new forecast that gets translated into a unique trading pattern. Those combination models do not explicitly consider investor preference. In fact, the combination methods are often based on some function of the forecast errors. The portfolio-based trading approach explicitly uses the parameters of the investor preference function. This is done by creating a subtle mix of multiple trading patterns. In that way, the portfolio-based trading approach customizes the portfolio to the investor’s specific preferences.

To obtain the “diversification” effect mentioned above, a four-component portfolio is constructed. Each component corresponds to one of the series of forecasts. Thus, the objective of the goal program is to determine the weights for allocating the investment to each component in the portfolio. Depending on the sign of the forecasts of the monthly excess return, a component will be put in either one of two positions. If the predicted excess return for the next month is positive, the investor will buy the matching index fund and short the one-month T-bill. Otherwise, if the predicted excess return for the next month is negative, the investor will instead short the index fund and long the one-month T-bill. This trading logic is based on the assumptions that an investor has no initial capital and needs to borrow money from the market and that he has to close his portfolio position at the end of each month. Let \( \hat{R}_{i,t+1} \) denote the excess return for next month (period \( t+1 \)) forecast by technique \( i \) (\( i = \text{NN}, \text{AES}, \text{ARIMA}, \text{MTF} \)). The trading rule for component \( i \) can be summarized as follows:

If \( (\hat{R}_{i,t+1} > 0) \) then

long index fund and short T-bill.

Else if \( (\hat{R}_{i,t+1} \leq 0) \) then

short index fund and long T-bill.

The composition of the portfolio is proportional to the weights determined by the GP3 model. In other words, component \( i \) of the portfolio is weighted by the solution \( \{w_i| i = \text{NN}, \text{AES}, \text{ARIMA}, \text{MTF}\} \) of GP3. Suppose the total principal for investment in the beginning of the month is fixed at \( M \) dollars. The asset allocation to component \( i \) of the portfolio is \( Mw_i \) such that

\[
\sum_i Mw_i = M. \tag{27}
\]

Based on this method of constructing the portfolio, the composite portfolio should exhibit the same historical distributional properties as the mean–variance–skewness efficient portfolio found by GP3. This approach to portfolio selection also allows the investor to diversify the risk due to the
variability of forecasts estimated by different techniques.

5. Performance evaluation and comparison

5.1. Experiment

The empirical experiment performed in this study contains three phases. In the first phase, we estimate the forecasts of a particular stock market index using the forecasting models outlined in Section 3. The first 228 observations (from January 1967 to December 1985) of the in-sample period are used as the training set to determine the model parameters and specifications. The prediction models are applied to forecast the index return for the rest of the in-sample period (from January 1986 to December 1990). As a result, four series of 60 in-sample forecasts, which represent the forecasts supplied by different analysts, are generated. Each individual forecasting model is then validated by its forecast performance on this in-sample validation period.

In the second phase, the monthly returns for each series of in-sample forecasts are measured. Based on these 60 observations, the distributional properties, i.e., mean, variance, and skewness, of the return on index are computed. Table 2 shows the average return on index trading using the forecasts supplied by each forecasting technique. Readers should be reminded that the returns are generated by a “no investment, no principal” trading scheme and thus cannot be directly compared to the on-going rates of return in the financial market. In addition to the first moment (average monthly return), second and third moments are also computed for the forecasts made within the in-sample period. Tables 3 and 4 show the estimated variance–covariance and skewness–coskewness matrices, respectively, for each financial index. Since the variance/covariance represent the square term while the skewness/coskewness is a cube term, the magnitudes of these descriptive statistics are quite small.

In addition to the statistical moments, the aspired levels $Z_1$, $Z_2$, and $Z_3$ are found by solving the subproblems GP2A, GP2B, and GP2C, respectively. The results of these GP subproblems are stated in Table 5. The aspired levels as well as the three moments are then entered to the multi-objective goal program GP3. The solution of GP3 indicates the weight set $W$ for constructing the composite portfolio explained in the last section. The procedure is repeated for different investor preferences. For exposition purpose, the optimal weight sets to construct the S&P 500, FTSE 100, and Nikkei 225 portfolios given an investor preference of $(p_1 = 1, p_2 = 1, p_3 = 1)$ are reported in Table 6. It should be noted that preference set $(1, 1, 1)$ was an extreme case where the weights for ARIMA were found to be zero. Nevertheless, the weights were strictly positive in other preference sets, implying that there is still an incremental benefit to adopt the forecasts of a below-average analyst.

In phase three, the four forecasting models are used to generate in the reserved out-of-sample period (from January 1991 to December 1995). Trading returns are computed accordingly. Using the weights estimated in phase two, we construct a composite portfolio and measure its returns. The entire phase is then repeated for different investor preferences and stock indices.

Table 2
Average return on index trading over the in-sample period from January 1986 to December 1991 (total of 60 observations)*

<table>
<thead>
<tr>
<th>Average rate of return</th>
<th>NN</th>
<th>AES</th>
<th>ARIMA</th>
<th>MTF</th>
</tr>
</thead>
<tbody>
<tr>
<td>US S&amp;P 500</td>
<td>0.004220</td>
<td>0.010585</td>
<td>0.012617</td>
<td>0.015197</td>
</tr>
<tr>
<td>UK FTSE 100</td>
<td>0.017984</td>
<td>0.011874</td>
<td>0.011599</td>
<td>0.012450</td>
</tr>
<tr>
<td>Japan Nikkei 225</td>
<td>0.013056</td>
<td>0.018280</td>
<td>0.018276</td>
<td>0.023505</td>
</tr>
</tbody>
</table>

*The four forecasting techniques used in our experiment are: neural network (NN), adaptive exponential smoothing (AES), ARIMA, and multivariate transfer function (MTF).
Table 3
Estimated variance–covariance matrices of the distributions of returns on index trading over the in-sample period from January 1986 to December 1991 (total of 60 observations)*

<table>
<thead>
<tr>
<th>Forecasting technique</th>
<th>Panel A: US S&amp;P 500</th>
<th>Panel B: UK FTSE 100</th>
<th>Panel C: Japan Nikkei 225</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NN $\sigma_{ij}$</td>
<td>AES $\sigma_{ij}$</td>
<td>ARIMA $\sigma_{ij}$</td>
</tr>
<tr>
<td>NN $\sigma_{ij}$</td>
<td>0.001544</td>
<td>0.000712</td>
<td>0.000781</td>
</tr>
<tr>
<td>AES $\sigma_{ij}$</td>
<td>0.000712</td>
<td>0.001450</td>
<td>0.001097</td>
</tr>
<tr>
<td>ARIMA $\sigma_{ij}$</td>
<td>0.000781</td>
<td>0.001097</td>
<td>0.001403</td>
</tr>
<tr>
<td>MTF $\sigma_{ij}$</td>
<td>0.000844</td>
<td>0.001129</td>
<td>0.001193</td>
</tr>
</tbody>
</table>

*The four forecasting techniques used in our experiment are: neural network (NN), adaptive exponential smoothing (AES), ARIMA, and multivariate transfer function (MTF). The highlighted values on the diagonals represent the variance of returns for the corresponding forecasting technique.

Table 4
Estimated skewness–coskewness matrices of the distributions of returns on index trading over the in-sample period from January 1986 to December 1991 (total of 60 observations)*

<table>
<thead>
<tr>
<th>Forecasting technique</th>
<th>Panel A: US S&amp;P 500</th>
<th>Panel B: UK FTSE 100</th>
<th>Panel C: Japan Nikkei 225</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NN $s_{ij}$</td>
<td>AES $s_{ij}$</td>
<td>ARIMA $s_{ij}$</td>
</tr>
<tr>
<td>NN $s_{ij}$</td>
<td>0.0000585</td>
<td>0.0000884</td>
<td>0.0000853</td>
</tr>
<tr>
<td>AES $s_{ij}$</td>
<td>0.0000565</td>
<td>0.0000637</td>
<td>0.0000687</td>
</tr>
<tr>
<td>ARIMA $s_{ij}$</td>
<td>0.0000518</td>
<td>0.0000668</td>
<td>0.0000565</td>
</tr>
<tr>
<td>MTF $s_{ij}$</td>
<td>0.0000459</td>
<td>0.0000601</td>
<td>0.0000556</td>
</tr>
</tbody>
</table>

*The four forecasting techniques used in our experiment are: neural network (NN), adaptive exponential smoothing (AES), ARIMA, and multivariate transfer function (MTF). The highlighted values on the diagonals represent the skewness of returns for the corresponding forecasting technique.
Table 5
Goal programming solutions of GP2 subproblems

<table>
<thead>
<tr>
<th></th>
<th>Subproblem GP2A</th>
<th>Subproblem GP2B</th>
<th>Subproblem GP2C</th>
</tr>
</thead>
<tbody>
<tr>
<td>US S&amp;P 500</td>
<td>0.015197</td>
<td>0.001057</td>
<td>7.15409 x 10^-3</td>
</tr>
<tr>
<td>UK FTSE 100</td>
<td>0.017964</td>
<td>0.001902</td>
<td>0.000264458</td>
</tr>
<tr>
<td>Japan Nikkei 225</td>
<td>0.023505</td>
<td>0.001627</td>
<td>2.29982 x 10^-3</td>
</tr>
</tbody>
</table>

Solving subproblems GP2A, GP2B, and GP2C yields the aspired levels, $Z_1^*$, $Z_2^*$, and $Z_3^*$, respectively. Each aspired level indicates the best case scenario for a particular objective without considering other objectives.

Table 6
Portfolio composition weights for the case of investor preference (1,1,1)

<table>
<thead>
<tr>
<th></th>
<th>$w_{NN}$</th>
<th>$w_{AES}$</th>
<th>$w_{ARIMA}$</th>
<th>$w_{MTF}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composition weights for S&amp;P 500 portfolio ($w_i$)</td>
<td>0.0000</td>
<td>0.3614</td>
<td>0.0000</td>
<td>0.6386</td>
</tr>
<tr>
<td>Composition weights for FTSE 100 portfolio ($w_i$)</td>
<td>0.5781</td>
<td>0.4219</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Composition weights for Nikkei 225 portfolio ($w_i$)</td>
<td>0.1046</td>
<td>0.3792</td>
<td>0.0000</td>
<td>0.5162</td>
</tr>
</tbody>
</table>

Each composition weight indicates the relative proportion of investment allocated to a particular component. There are four trading components in the index portfolio, each guided by the forecasts estimated by one of the four forecasting models, namely, neural network (NN), adaptive exponential smoothing (AES), ARIMA, and multivariate transfer function (MTF). The results presented here are based on an investor preference of (1,1,1), i.e., equal preference over average, variance, and skewness of returns.

5.2. Performance measure

It should be noted that a direct comparison between the results from the multiobjective model and those generated by a single-objective model is not practical. This is because the tradeoffs inherent to the multiobjective optimization leave the approach some disadvantages with respect to the criterion a single-objective model is based on. On the other hand, the conflicting goals in the multiobjective model cannot be captured by those models which solely rely on single objectives. Given this notion, we use the following preference function to evaluate the performance of the various models:

$$Z = \left| \frac{Z_1^* - Z_1}{Z_1^*} \right|^{p_1} + \left| \frac{Z_2^* - Z_2}{Z_2^*} \right|^{p_2} + \left| \frac{Z_3^* - Z_3}{Z_3^*} \right|^{p_3}.$$

This function is essentially a variation of the objective function (26a) found in GP3 formulation. $Z_1^*$, $Z_2^*$, and $Z_3^*$ indicate the best average, variance, and skewness, respectively, of the out-of-sample returns among all tested models. Similarly, $Z_1$, $Z_2$, and $Z_3$, are the average, variance, and skewness of the returns for a particular model while $Z$ is the aggregate performance measure adjusted by the investor preference ($p_1$, $p_2$, $p_3$). The normalization allows us to directly compare and add the values of the objectives, thus avoiding the problem of bias found in Lai (1991) and Chunhachinda et al. (1997).

To verify the robustness and sensitivity of the proposed multiobjective approach to changes in the investor’s preference, different levels of preference are investigated. Specifically, preferences of (1,1,1), (1,2,2), (2,2,1), (1,2,1), and (1,1,0) are included in our experiment. The results based on the (1,1,1) preference structure means that mean, variance, and skewness of return are of equal importance to the investor while those of (1,2,2) and (1,2,1) give more emphasis to risk control. (1,1,0) is a benchmark case, representing the Markowitz mean–variance portfolio.

5.3. Comparative evaluation

In order to draw robust conclusions, the evaluation should involve statistical comparison with other models. The extensive empirical evaluation by Russel and Adam (1987) provides a list of models to combine forecasts. Among the candidates, our pilot study suggests a group of models:

- **CSAVA** – simple average weighting of all models,
- **CMAD** – selective weighting based upon inverse proportion to MAE,
CMSE – selective weighting based upon inverse proportion to MSE,
CCIV – selective weighting based upon absolute error performance.

Furthermore, the four individual forecasting models used in the portfolio approach are also included in our empirical evaluation. This can verify the efficacy of the proposed approach over those simpler models driven by a single forecasting technique. Hence, returns on trading guided by individual forecasting techniques and a number of forecast-combining models are computed.

For each index and preference level, the experimental procedure described above yields only a single value of aggregate measure $Z$. Thus, bootstrap re-sampling is performed to ensure the obtained results are not by chance. The design of our bootstrap re-sampling plan involves the generation of 1500 bootstrap samples for each scenario. The mean and standard error for each group of 1500 samples thus provide a basis for $t$-tests. The experimental procedure for bootstrap re-sampling can be found in the Appendix. Interested readers can also refer to Efron (1990) and Shao and Tu (1995) for a more detailed explanation.

6. Results and discussions

Table 7 tabulates the aggregate performance $Z$ of the proposed composite portfolio and various forecasting models over the out-of-sample period. The reported mean and standard error (SE) of $Z$ are based on 1500 bootstrap samples. It is noteworthy to point out that the sample means and standard errors cannot be compared across different preference levels because the magnitude of values is relative to a specific preference level. This notion is true especially for comparison of the case of $(1,1,0)$, which represents a Markowitz mean–variance efficient portfolio, with other models which account for mean, variance, and skewness. A $t$-test for the difference of means between a particular model (i.e., one of the forecasting-combining models or the individual forecasting techniques stated in Section 5.3) and the portfolio is conducted. Results show that the $Z$ value for the portfolio is significantly better than any other examined model at $\alpha = 0.05$ level. This conclusion is valid for all indices and investor preferences considered in our experiment.

An explanation of the results may offer some insights to identify the characteristics of our proposed approach and possibly some guidance to better utilize/apply the research. The portfolio yields significantly better performance than those single forecasting techniques (NN, AES, ARIMA, MTF). An explanation for this is that single forecasting technique, unlike the portfolio, is unable to benefit from diversification. As a result, the overall performance of trading which is solely driven by a particular forecasting technique can be dampened by unexpected stock market movements, causing inconsistency in performance throughout the time. To alleviate this problem, some studies in the literature suggest using a combination of forecasts. These studies find out that the predictability improves when a number of forecasts are combined. However, an examination of the performance of these models (CSAVA, CMAD, CMSE, CCIV) relative to the portfolio shows that they are as weak as the single forecasting techniques. It is possibly because the combination models do not explicitly account for the risk factors (variance and skewness) inherent to investor’s preference. Also, the forecasts made by most combination models are generated by a single objective of minimization of some function of forecast error. This approach seems to be less effective than the multiobjective portfolio approach which explicitly considers the investment return.

A closer look at Table 7 suggests that the advantage of the portfolio-based approach over the other models is most apparent for the Japanese Index. This can be seen by looking at the $t$-statistics: those related to the Japanese Nikkei are consistently higher. The $t$-statistics are the valid measures to compare across indices and models because they are scale free. They measure the standardized deviation of the performance of the investigated models from that of the portfolio-based approach. It is suspected that the higher relative benefit from the adoption of portfolio-based approach in Nikkei trading is due to the
Table 7
Relative performance of the composite portfolio and other forecasting models over the out-of-sample period from January 1991 to December 1995 (total of 60 observations)*

<table>
<thead>
<tr>
<th></th>
<th>US S&amp;P 500</th>
<th></th>
<th>UK FTSE 100</th>
<th></th>
<th>Japan Nikkei 225</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SE</td>
<td>r Stat</td>
<td>Mean</td>
<td>SE</td>
<td>r Stat</td>
</tr>
<tr>
<td><strong>Value of Z for investor preference (1, 1, 1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSAVA</td>
<td>1.6926</td>
<td>0.0177</td>
<td>2.45b</td>
<td>3.6347</td>
<td>0.2388</td>
<td>10.69b</td>
</tr>
<tr>
<td>CMAD</td>
<td>1.9517</td>
<td>0.0184</td>
<td>16.41b</td>
<td>3.6418</td>
<td>0.2436</td>
<td>10.50b</td>
</tr>
<tr>
<td>CMSE</td>
<td>1.9403</td>
<td>0.0183</td>
<td>15.96b</td>
<td>3.9173</td>
<td>0.2839</td>
<td>9.98b</td>
</tr>
<tr>
<td>CCIv</td>
<td>1.9464</td>
<td>0.0181</td>
<td>16.45b</td>
<td>3.6926</td>
<td>0.1655</td>
<td>15.77b</td>
</tr>
<tr>
<td>NN</td>
<td>2.0955</td>
<td>0.0227</td>
<td>19.66b</td>
<td>1.8608</td>
<td>0.1046</td>
<td>7.44b</td>
</tr>
<tr>
<td>AES</td>
<td>1.9288</td>
<td>0.0184</td>
<td>15.19b</td>
<td>3.8975</td>
<td>0.3301</td>
<td>8.53b</td>
</tr>
<tr>
<td>ARIMA</td>
<td>3.1560</td>
<td>0.0386</td>
<td>39.06b</td>
<td>4.1023</td>
<td>0.0869</td>
<td>34.77b</td>
</tr>
<tr>
<td>MTF</td>
<td>2.0267</td>
<td>0.0184</td>
<td>20.50b</td>
<td>4.7794</td>
<td>0.4028</td>
<td>9.18b</td>
</tr>
<tr>
<td>Portfolio</td>
<td>1.6491</td>
<td>0.0157</td>
<td></td>
<td>1.0824</td>
<td>0.0475</td>
<td></td>
</tr>
<tr>
<td><strong>Value of Z for investor preference (1, 2, 2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSAVA</td>
<td>253.40</td>
<td>12.1587</td>
<td>13.50b</td>
<td>642.08</td>
<td>83.9938</td>
<td>6.02b</td>
</tr>
<tr>
<td>CMAD</td>
<td>282.83</td>
<td>10.4757</td>
<td>18.48b</td>
<td>778.45</td>
<td>84.5982</td>
<td>7.59b</td>
</tr>
<tr>
<td>CMSE</td>
<td>285.85</td>
<td>10.6479</td>
<td>18.46b</td>
<td>908.89</td>
<td>97.8977</td>
<td>7.89b</td>
</tr>
<tr>
<td>CCIv</td>
<td>280.99</td>
<td>10.6543</td>
<td>18.00b</td>
<td>905.17</td>
<td>100.8420</td>
<td>7.63b</td>
</tr>
<tr>
<td>NN</td>
<td>332.83</td>
<td>12.9730</td>
<td>18.78b</td>
<td>270.79</td>
<td>62.9585</td>
<td>2.14b</td>
</tr>
<tr>
<td>AES</td>
<td>285.57</td>
<td>11.0996</td>
<td>17.69b</td>
<td>1003.91</td>
<td>101.3165</td>
<td>8.57b</td>
</tr>
<tr>
<td>ARIMA</td>
<td>484.87</td>
<td>14.8099</td>
<td>26.7b</td>
<td>1169.05</td>
<td>105.1601</td>
<td>9.82b</td>
</tr>
<tr>
<td>MTF</td>
<td>297.03</td>
<td>10.7701</td>
<td>19.29b</td>
<td>1380.81</td>
<td>111.2475</td>
<td>11.19b</td>
</tr>
<tr>
<td>Portfolio</td>
<td>89.24</td>
<td>3.3397</td>
<td></td>
<td>136.04</td>
<td>54.6553</td>
<td></td>
</tr>
<tr>
<td><strong>Value of Z for investor preference (2, 2, 1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSAVA</td>
<td>146.79</td>
<td>5.3158</td>
<td>15.72b</td>
<td>182.19</td>
<td>3.6489</td>
<td>39.55b</td>
</tr>
<tr>
<td>CMAD</td>
<td>189.51</td>
<td>5.8338</td>
<td>21.65b</td>
<td>133.29</td>
<td>3.5374</td>
<td>26.97b</td>
</tr>
<tr>
<td>CMSE</td>
<td>190.17</td>
<td>5.9344</td>
<td>21.39b</td>
<td>138.48</td>
<td>3.5914</td>
<td>27.97b</td>
</tr>
<tr>
<td>CCIv</td>
<td>188.46</td>
<td>5.6027</td>
<td>22.35b</td>
<td>130.48</td>
<td>3.3108</td>
<td>27.97b</td>
</tr>
<tr>
<td>NN</td>
<td>181.14</td>
<td>6.3877</td>
<td>18.46b</td>
<td>93.05</td>
<td>2.6962</td>
<td>20.46b</td>
</tr>
<tr>
<td>AES</td>
<td>189.21</td>
<td>6.0894</td>
<td>20.69b</td>
<td>111.33</td>
<td>3.1128</td>
<td>23.60b</td>
</tr>
<tr>
<td>ARIMA</td>
<td>249.67</td>
<td>7.7962</td>
<td>23.91b</td>
<td>136.10</td>
<td>3.3928</td>
<td>28.95b</td>
</tr>
<tr>
<td>MTF</td>
<td>200.26</td>
<td>6.0321</td>
<td>22.72b</td>
<td>128.01</td>
<td>3.4263</td>
<td>26.31b</td>
</tr>
<tr>
<td>Portfolio</td>
<td>63.23</td>
<td>2.0141</td>
<td></td>
<td>37.88</td>
<td>1.4749</td>
<td></td>
</tr>
<tr>
<td><strong>Value of Z for investor preference (1, 2, 1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSAVA</td>
<td>239.73</td>
<td>11.2767</td>
<td>17.65b</td>
<td>103.27</td>
<td>3.5856</td>
<td>28.06b</td>
</tr>
<tr>
<td>CMAD</td>
<td>282.75</td>
<td>12.4805</td>
<td>19.40b</td>
<td>97.10</td>
<td>3.7680</td>
<td>25.06b</td>
</tr>
<tr>
<td>CMSE</td>
<td>285.32</td>
<td>12.3336</td>
<td>19.84b</td>
<td>103.07</td>
<td>3.8726</td>
<td>25.93b</td>
</tr>
<tr>
<td>CCIv</td>
<td>281.20</td>
<td>12.8198</td>
<td>18.76b</td>
<td>96.67</td>
<td>3.5105</td>
<td>26.78b</td>
</tr>
<tr>
<td>NN</td>
<td>296.48</td>
<td>15.2449</td>
<td>18.76b</td>
<td>87.62</td>
<td>3.3460</td>
<td>25.39b</td>
</tr>
<tr>
<td>AES</td>
<td>281.77</td>
<td>12.4025</td>
<td>19.44b</td>
<td>94.55</td>
<td>3.3941</td>
<td>27.07b</td>
</tr>
<tr>
<td>ARIMA</td>
<td>308.77</td>
<td>13.3735</td>
<td>20.05b</td>
<td>99.16</td>
<td>3.6146</td>
<td>26.70b</td>
</tr>
<tr>
<td>MTF</td>
<td>292.95</td>
<td>12.7814</td>
<td>19.74b</td>
<td>100.08</td>
<td>3.6817</td>
<td>26.46b</td>
</tr>
<tr>
<td>Portfolio</td>
<td>40.67</td>
<td>3.5820</td>
<td></td>
<td>2.66</td>
<td>0.2678</td>
<td></td>
</tr>
<tr>
<td><strong>Value of Z for investor preference (1, 1, 0)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSAVA</td>
<td>1.1429</td>
<td>0.0195</td>
<td>10.60b</td>
<td>1.7995</td>
<td>0.0200</td>
<td>63.52b</td>
</tr>
<tr>
<td>CMAD</td>
<td>1.4490</td>
<td>0.0196</td>
<td>26.11b</td>
<td>1.4820</td>
<td>0.0207</td>
<td>45.92b</td>
</tr>
<tr>
<td>CMSE</td>
<td>1.4455</td>
<td>0.0195</td>
<td>26.06b</td>
<td>1.5166</td>
<td>0.0208</td>
<td>47.48b</td>
</tr>
<tr>
<td>CCIv</td>
<td>1.4445</td>
<td>0.0193</td>
<td>26.38b</td>
<td>1.4766</td>
<td>0.0204</td>
<td>46.45b</td>
</tr>
<tr>
<td>NN</td>
<td>1.3116</td>
<td>0.0225</td>
<td>16.67b</td>
<td>1.1947</td>
<td>0.0177</td>
<td>37.53b</td>
</tr>
</tbody>
</table>

*Significance levels: 5% (b), 2% (a), and 1% ( serious).
unstable and highly volatile environment of the Japanese market during our forecast period (1990–1995). With this notion, we believe that our portfolio-based approach can derive more benefit relative to other methods when the investment environment becomes more volatile and unstable. This is because the portfolio is able to explicitly take into account the risk factors and control them in a way that best fulfills the investor’s preference. In other words, the portfolio-based approach is able to reduce risk (by reducing the variance) and hedge against the worst case scenarios (by increasing the skewness). These observations lead us to think that the portfolio-based approach has a bright future in applications to the more volatile and fast expanding emerging markets. This perspective is particularly exciting, if we keep in mind that it is in those emerging markets that will arise most of the opportunities of the next millennium. However, only those investors that are best equipped to avoid the dangers will profit from those opportunities. The portfolio-based approach seems like a powerful tool in that respect.

Because of the empirical nature of our experiment, we would like to note that the use of any data set for forecasting purposes incurs the risks of data-dependent and estimated models not necessarily retain their performance characteristics in the future. Given that the real success in any forecasting application will only come in the future when the forecasting model under consideration is really tested, this study attempts to mitigate model selection/data-dependency risks by carefully partitioning the data in such a way that the model selection and specification process do not make use of any information from the out-of-sample evaluation period. This mimics the information set available to a real-world forecaster. Also, in this study, the basic structure and initial formulation of our forecasting models were derived from economic arguments rather than from statistical extrapolation alone. This also helps to reduce the risks associated with the use of data for forecasting. Finally, a process that stays relatively unchanged will have a better chance at being forecast correctly.

7. Conclusions

This paper proposes a multiobjective approach to combine the forecasts obtained from different analysts (or sources). The approach examines the historical performance of the various series of

Table 7 (Continued)

<table>
<thead>
<tr>
<th></th>
<th>US S&amp;P 500</th>
<th></th>
<th>UK FTSE 100</th>
<th></th>
<th>Japan Nikkei 225</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SE</td>
<td>t Stat</td>
<td>Mean  SE</td>
<td>t Stat</td>
<td>Mean  SE</td>
</tr>
<tr>
<td>AES</td>
<td>1.4243 0.0200</td>
<td>24.46a</td>
<td>1.3207 0.0205</td>
<td>38.58a</td>
<td>1.5661 0.0217</td>
</tr>
<tr>
<td>ARIMA</td>
<td>1.7855 0.0275</td>
<td>30.89b</td>
<td>1.5080 0.0205</td>
<td>47.69b</td>
<td>1.7387 0.0223</td>
</tr>
<tr>
<td>MTF</td>
<td>1.4947 0.0200</td>
<td>27.88b</td>
<td>1.4387 0.0212</td>
<td>42.88b</td>
<td>1.7726 0.0201</td>
</tr>
<tr>
<td>Portfolio</td>
<td>0.93616 0.0144</td>
<td></td>
<td>0.5296 0.0121</td>
<td></td>
<td>0.4492 0.0098</td>
</tr>
</tbody>
</table>

a The aggregate performance measure Z is calculated by Eq. (28):

\[
Z = \left| \frac{Z_1^* - Z_1}{Z_1^*} \right|^p + \left| \frac{Z_2^* - Z_2}{Z_2^*} \right|^p + \left| \frac{Z_3^* - Z_3}{Z_3^*} \right|^p
\]

Z_1^*, Z_2^*, and Z_3^* indicate the best average, variance, and skewness, respectively, of the out-of-sample returns among all tested models. Similarly, Z_1, Z_2, and Z_3, are the average, variance, and skewness of the returns for a particular model while Z is the aggregate performance measure adjusted by the investor preference (p_1, p_2, p_3). The first, second and third terms measure the performance relative to the best scenarios of average, variance, and skewness of returns, respectively. All terms are normalized and thus additive.

b Next to the t statistic indicates that the mean of Z over 1500 bootstrap samples for a particular forecasting technique is significantly different from that of the composite index portfolio at α = 0.05 level. “SE” stands for standard error of the bootstrap samples.

CSAVA – simple average weighting of all models; CMAD – selective weighting based upon inverse proportion to MAE; CMSE – selective weighting based upon inverse proportion to MSE; CCTV – selective weighting based upon absolute error performance.

There are four components in the portfolio, each guided by the forecasts estimated by one of the four forecasting techniques, i.e., neural network (NN), adaptive exponential smoothing (AES), ARIMA, and multivariate transfer function (MTF).
forecasts and combines them based on the average, variance, and skewness of investment returns. Through the use of a GP model, an investor can construct a portfolio which matches his or her preference. This portfolio-based approach also adds the benefits of diversification in trading. We test our proposed approach with three widely traded broad market indices, S&P 500, FTSE 100, and Nikkei 225. Improved performance of the multiobjective portfolio approach relative to those of individual forecasting techniques and some previously suggested forecast-combining models is measured. The empirical results indicate that the performance of the proposed approach statistically outperforms the others at a significance level of 0.05. Moreover, we find that the benefits of our approach become more apparent when the market exhibits higher volatility and instability. This makes the approach more meaningful to those investors in the emerging markets. As a final word, more indices, especially the ones of emerging markets, should be examined in future research. A similar research avenue can be applied to other financial instruments such as individual stocks and foreign currencies.

Acknowledgements

The authors wish to thank Doug Blocher, Sreenivas Kamma, Robert Klemkosky, Roger Schmenner, Munirpallam Venkataramanan, and Wayne Winston for their valuable inputs and comments. The errors in this paper are our own.

Appendix A. Economic rationale for macroeconomic state variables

In the light of the previous literature, it is hypothesized that various measures of the macroeconomic environment may be used as input variables in the construction of prediction models to forecast the stock market index. Table 1 outlines an array of such macroeconomic state variables which are applied to the paper. We will continue to describe some of the economic intuition concerning why the state variables chosen in this study are expected to indicate future stock market movement.

The term structure of interest rate (TS), that is, the spreads of long-term bond yields (LT) over short-term bond yields (ST) may have some power to forecast stock returns. The hypothesis that this variable may have some power in forecasting stock returns is supported by the observation that this variable has a business cycle pattern. It is low around business peaks and high around business troughs. Thus, the term structure of interest rate captures the cyclical variation in expected returns. This fact, combined with the historical evidence which shows that stock returns are generally lower during recessions, substantiates the notion that term spread may exhibit some degree of predictive power on stock returns. This is because a large term spread may suggest probable business expansion or increased economic activity in the future that corresponds to higher stock returns. In short, the term spread variable may be thought of as an indicator of the future level of economic activity which then, indirectly, result in some power to forecast stock returns.

Short-term interest rates (ST) also fluctuate with economic conditions. T-bill rates tend to be low in a business contraction, especially at the low turning points of business cycles. Therefore low T-bill rates may indicate the future business expansion or increased economic activity to certain extent. Business expansions or increased economic activity has been historically associated with higher stock returns and recessions with lower stock returns. Like the term structure variable, the short-term interest rate may also be thought of as an indicator of the future level of economic activity which then, indirectly, result in some power to forecast stock returns.

The lagged index return \( R \) is included in this study to check whether the time series properties of the past index returns contain any information that is useful in forecasting the future index returns. Macroeconomic state variables CPI and IP are also included in our examination as they possibly contain imperative information concerning the forecast of future stock index returns. Whether
these variables are positively or negatively correlated with future stock index returns is uncertain. To be specific, if CPI and IP turn out to be reasonable proxies for the current health of the economy, then they will be negatively correlated with future index returns. On the other hand, if these variables turn out to be reasonable proxies for the future growth rates of the economy, they will be positively correlated with future index returns. In general, the impact of CPI and IP on future index return depends on whether these variables proxy the current health of the economy or the future growth rates of the economy.

Appendix B. Bootstrap re-sampling

The bootstrap technique is aimed at deriving the distribution of a statistic from a sample of data. The idea is that a sample of data from a population will provide one value of the statistic of interest. However, it is often the case that we are interested in knowing the distribution of that statistic. This is essential in a case where we want to test hypotheses about the statistic of interest. In fact, hypothesis testing often requires that the variance of the statistic be known. Then the problem becomes that of inferring the distribution of a statistic from one observation. The solution to this problem is bootstrapping.

Bootstrapping is based on two basic assumptions:

1. The sample of data is generated from a multinomial distribution with possible values similar to those in the sample.
2. The data generated from the distribution are identically and independently distributed (IID). Given those two assumptions, the bootstrap methodology will derive the empirical distribution of the statistic from the sample of data as follows:

   1. Generate \( N \) number of new samples of data based on re-sampling from the original data sample with replacement. The idea is that the original sample is a best “guess” at the “true” distribution of the data generating process. Resampling with replacement from the original sample mimics drawing IID observations randomly from the hypothesized true distribution of the data.

   2. Compute the statistic of interest from each of the \( N \) “bootstrapped” samples. This will give us \( N \) values of the statistics. Those \( N \) values will represent the empirical distribution of the statistic. Now that we have the distribution, we are able to compute its moments (mean, variance...) and use them in hypothesis testing.

   In our case, the original sample of data is the series of returns. We generate \( N = 1500 \) bootstrap samples. The statistic of interest is the aggregate performance measure \( Z \). We are able to obtain 1500 \( Z \)’s which allows us to compute the mean and variance of the distribution of \( Z \) for all the forecasting methodologies used in the paper. From this point, we are able to test hypothesis concerning the difference between the \( Z \) of each technique and that of the portfolio formed by GP.

References


