PERGAMON

Computers & Operations Research 30 (2003) 901–923

computers & operations research

www.elsevier.com/locate/dsw

Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index

An-Sing Chen[a,*], Mark T. Leung[b], Hazem Daouk[c]

[a] Department of Finance, National Chung Cheng University, Ming-Hsiung, Chia-Yi 621, Taiwan
[b] Department of Management Science and Statistics, College of Business, University of Texas, San Antonio, TX 78249, USA
[c] Department of Applied Economics, Cornell University, Ithaca, NY 14853, USA

Abstract

In this study, we attempt to model and predict the direction of return on market index of the Taiwan Stock Exchange, one of the fastest growing financial exchanges in developing Asian countries. Our motivation is based on the notion that trading strategies guided by forecasts of the direction of price movement may be more effective and lead to higher profits. The probabilistic neural network (PNN) is used to forecast the direction of index return after it is trained by historical data. Statistical performance of the PNN forecasts are measured and compared with that of the generalized methods of moments (GMM) with Kalman filter. Moreover, the forecasts are applied to various index trading strategies, of which the performances are compared with those generated by the buy-and-hold strategy as well as the investment strategies guided by forecasts estimated by the random walk model and the parametric GMM models. Empirical results show that the PNN-based investment strategies obtain higher returns than other investment strategies examined in this study. Influences of length of investment horizon and commission rate are also considered.

Keywords: Emerging economy; Forecasting; Trading strategy; Neural networks; Generalized methods of moments (GMM)

1. Introduction

Although there exists some studies which deal with the issues of forecasting stock market index and development of trading strategies, most of the empirical findings are associated with the developed

*Corresponding author. Tel.: 011-886-5-272-0411x34201; fax: 011-886-5-272-0818.
E-mail addresses: finasc@ccunix.ccu.edu.tw (A.-S. Chen), mtleung@utsa.edu (M.T. Leung), hdaouk@umich.edu (H. Daouk).

PII: S0305-0548(02)00037-0

financial markets (e.g., US, UK, and Japan). Nowadays, many international investment bankers and brokerage firms have major stakes in overseas markets. Given the economic success of Taiwan in the last two decades, the financial markets in this Asian country have attracted considerable global investments. This view is further corroborated by the recent introduction of several Taiwan Stock Index instruments by Singapore International Monetary Exchange (SIMEX) in January 1997. Realizing the growing importance of the Taiwanese stock market and its influence on the current Asian financial crisis, our study attempts to develop effective forecasting models for predicting the Taiwan Stock Index returns. There are two basic reasons for a closer examination of this index trading vehicle. First, it provides an effective means for the investors to hedge against potential market risks. Second, it creates new profit making opportunities for market speculators and arbitrageurs. Therefore, being able to accurately forecast stock market index has profound implications and significance to researchers and practitioners alike.

Another motivation for this study is to confirm whether we can extend some basic notions of traditional financial forecasting modeling, which are built upon the observation of well established financial systems, to a rapidly growing emerging economy. Champion [1] presents two diametrically opposed views of the Taiwanese market which differentiate it from the more developed financial markets. The author argues that “this market was different and had an internal logic of its own which allowed it to defy laws applicable elsewhere”. However, there is an opposing view such that “normal financial relationships must sooner or later prevail”. Therefore, given the rising popularity of index trading, it is of practical interest to assess the predictive strength of those explanatory variables, which are found to be useful in the forecasting of well established markets, in Taiwan stock market.

Our study models and predicts the Taiwan Stock Exchange (TSE) Index using neural networks. Their performance is compared with that of some parametric forecasting approaches such as generalized methods of moments (GMM) and random walk. To provide a more complete evaluation of the models, our comparison is based on not only the performance statistics but also the trading profits. Thus, this study develops a set of trading strategies to translate the forecasts into monetary returns. In addition, the experimental analysis investigates whether the length of the investment horizon has a significant impact on the quality of the forecasts. The remaining portion of this paper is organized as follows: A literature review and economic justification are given in the next section. In Section 3, we provide a description and conceptual foundation of the forecasting approaches (models) used in this study. Then, the results of forecasting are presented and discussed in Section 4. Section 5 describes the proposed index trading strategies which are driven by the forecasts made by various forecasting models. The last section concludes the paper.

2. Background

A lot of paper prove that it's valid to forecast stock using neural network in the well-established markets like US and Europe, but there are also some evidence showing that some emerging stock markets can be forecasted as well

2.1. Evidence of return predictability

There exists considerable evidence showing that stock returns are to some extent predictable. Most of the research is conducted using data from well established stock markets such as the US, Western Europe, and Japan. It is, thus, of interest to study the extent of stock market predictability using data from less well established stock markets such as that of Taiwan.

For the US, several studies examine the cross-sectional relationship between stock returns and fundamental variables. Variables such as earnings yield, cash flow yield, book-to-market ratio, and size are shown to have some power predicting stock returns. Banz and Breen [2], Jaffee et al. [3], and Fama and French [4] are good examples of this group of research. Further, studies based on European markets report similar findings. The results of Ferson and Harvey [5] indicate that returns are, to a certain extent, predictable across a number of European markets (e.g., UK, France, Germany). In their study which is aimed at forecasting the UK stock prices, Jung and Boyd [6] report "reasonably good" performance of their forecasts, suggesting that the predictive strength of their stock return models are not negligible. For the Japanese stock market, the empirical investigations by Jaffe and Westerfield [7] and Kato et al. [8] also find some evidence of predictability in the behavior of index returns.

Using time-series analysis, Fama and French [9] identify three common risk factors, an overall market factor, and some factors related to firm size and book-to-market equity which seem to explain the average returns on stocks and bonds. Moreover, Fama and Schwert [10], Rozeff [11], Keim and Stambaugh [12], Campbell [13], Fama and Bliss [14], and Fama and French [15–17] find out that macroeconomic variables such as short-term interest rates, expected inflation, dividend yields, yield spreads between long- and short-term government bonds, yield spreads between low- and high-grade bonds, lagged price–earnings ratios, and lagged returns have some power to predict stock returns. At the same time, the studies by Chen et al. [18] and Chan et al. [19] suggest that changes in aggregate production, inflation, the short-term interest rates, the slope of term structure (measured by the difference in returns on long- and short-term government bonds) and the risk premium (measured by the difference in returns on low- and high-grade bonds) are other macroeconomic factors that have some power to predict stock returns.

Although most of the papers in this avenue of research are related to the financial markets in developed economies, several recent articles do show that return predictability also exists in those less-developed financial markets. Ferson and Harvey [5] examine 18 international equity markets, some of which are found in developing economies. The study provides evidence of returns predictability. Harvey [20] focuses on emerging markets by looking at the returns of more than 800 equities from 20 emerging markets including Taiwan. He finds that the degree of predictability in the emerging markets is greater than that found in the developed markets. In addition, local information plays a much more important role in predicting returns in the emerging markets than in the developed markets. This characteristic helps explaining the difference in predictability between the two kinds of markets.

2.2. *Economic rationale*

Why we use some macro variables as independent variables

In light of the previous literature, it is hypothesized that various measures of the macroeconomic environment which is available to the forecaster may be used as input state variables in the construction of prediction models to forecast the direction of movement of the stock market index. Table 1 outlines an array of such macroeconomic state variables which are applied to the paper. In the following, we will describe 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), i.e., the spreads of long-term bond yields over short-term bond yields, may have some power to forecast stock returns. The hypothesis that this variable may

Table 1
List of potential economic state input variables and forecasted output variables

Input variables
TS—term structure proxy
Three-year government bond rate minus the 1-month risk-free rate. The 1-month risk-free rate is the 1-month deposit rate at the First Commercial Bank.
TB—short-term interest rate
One-month deposit rate at the First Commercial Bank.
DS3, *DS6*, *DS12—lagged index returns*
Continuously compounded lagged 3-, 6-, and 12-month annualized excess returns of the Taiwan index respectively.
GC3, *GC6*, *GC12*, *PC3*, *PC6*, *PC12—consumption level*
Continuously compounded lagged 3-, 6-, and 12-month annualized growth rates of government consumption and private consumption previous to the period being forecasted.
GNP3, *GNP6*, *GNP12*, *GDP3*, *GDP6*, *GDP12—gross national and domestic products*
Continuously compounded lagged 3-, 6-, and 12-month annualized growth rates of the gross national product and gross domestic product previous to the period being forecasted.
CPI3, *CPI6*, *CPI12*, *IP3*, *IP6*, *IP12—consumer price and production level*
Continuously compounded lagged 3-, 6-, and 12-month annualized growth rates of the consumer price index and industrial production previous to the period being forecasted.

Output variables
MR3, *MR6*, *MR12—returns on index*
Continuously compounded 3-, 6-, and 12-month annualized excess returns of the Taiwan 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.

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 also fluctuate with economic conditions. T-bill rates (TB) 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 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. The variables GC, PC, GNP, GDP, 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 state variables are positively or negatively correlated with future stock index returns is uncertain. To be specific, if these variables 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 these state variables on future index return depends on whether these variables proxy the current health of the economy or the future growth rates of the economy.

2.3. Forecasting the direction of index return

We forecast the direction of index return

Most trading practices adopted by financial analysts rely on accurate prediction of the price levels of financial instruments. However, some recent studies have suggested that trading strategies guided by forecasts on the direction of price change may be more effective and generate higher profits. Wu and Zhang [21] investigate the predictability of the direction of change in the future spot exchange rate. In another study, Aggarwal and Demaskey [22] provide evidence that the performance of cross-hedging improves significantly if the direction of changes in exchange rates can be predicted. Based on the S&P 500 futures, Maberly [23] explores the relationship between the direction of interday and intraday price change. O'Connor et al. [24] conduct a laboratory-based experiment and conclude that individuals are showing different tendencies and behaviors for upward and downward series. Finally, in their study on the All Ordinaries Index futures traded at the Australian Associated Stock Exchanges, Hodgson and Nicholls [25] suggest to hold an evaluation of the economic significance of the direction of price changes in future research. In summary, the findings in these studies are reasonable because an accurate point estimation, as judged by its deviation from the actual observation, may not be a good predictor of the direction of change in the instrument's price level. Also, predicting the direction is a practical issue which usually affects a financial trader's decision to buy or sell an instrument.

3. Predicting returns on Taiwan Stock Index

3.1. Data

Where the data from?

The data used in this study are obtained from the E.P.S. database maintained by the Department of Education of Taiwan. The data set covers the horizon from January 1982 to August 1992 and is divided into two periods: the first period runs from January 1982 to August 1987 and the second period runs from September 1987 to August 1992. The first period, the in-sample estimation period, is used for model selection and validation. The second period is the reserved out-of-sample evaluation period and is used to compare the forecasts and trading performances of various models. Depending on the length of investment horizon, the forecasted variable is the continuously compounded 3-month (MR3), 6-month (MR6), or 12-month (MR12) excess returns on the Taiwan index. The independent variables (see Table 1) for predicting the index returns are all observable on or before the last day of the month preceding the month corresponding to the first day of the forecast period. For instance, for the prediction of the 6-month ahead continuously compounded return starting on March 1, 1984, all independent variables must be observable on or before the last day of February 1984. Constructing the data set in this manner ensures that the generation of out-of-sample forecasts will be similar to those made in the real world. It is because only observable, but not future unobservable, data can be used as inputs to the forecasting models.

3.2. Neural network forecasting

Introduction to the Neural Network

Both academic researchers and practitioners have made tremendous efforts to predict the future movements of stock market and devise financial trading strategies to translate the forecasts into profits. Recently, in addition to econometric forecasting approaches, artificial neural networks (ANN) have been demonstrated to provide promising results in financial forecasting and trading. A comprehensive review of the fundamental concepts and principals of the ANN can be found in Rumelhart and McClelland [26] and Caudill and Butler [27]. Morever, Hawley et al. [28] and Medsker et al. [29] provide an overview of the neural network models in the fields of finance and investment.

3.2.1. Probabilistic neural network

The reason for using Neural Network

The neural network models used in this study are based on the topology of probabilistic neural network (PNN) proposed by Specht [30,31]. Technically, PNN is a classifier and is able to deduce the class/group of a given input vector after the training process is completed. There are a number of appealing features which justify our adoption of this type of neural network to this study. First, training of PNN is rapid, enabling us to develop a frequently updated training scheme. Essentially, the network is re-trained each time the data set is updated and thus the most current information can be reflected in estimation. Second, the logic of PNN is able to extenuate the effects of outliers and questionable data points and thereby reduces extra effort on scrutinizing training data. Third, and the most important, PNN provides the Bayesian probability of the class affiliation. The proposed trading strategies subsequently use this valuable information to make periodic decisions on asset allocation. The actual implementation will be discussed later in detail.

3.2.2. PNN logic

The logic of Neural Network

PNN is conceptually built on the Bayesian method of classification which, given enough data, is capable of classifying a sample with the maximum probability of success [32]. The principle of a Bayesian classifier rests on the selection of class i with the largest product term in the Bayesian Classification Theorem:

$$\max_i \{h_i l_i f_i(\boldsymbol{X})\}, \tag{1}$$

where h_i is the a priori probability for class i, l_i the loss incurred by misclassifying a sample which truly belongs to class i, $\boldsymbol{X}$ is $(x_1, x_2, \ldots, x_k)$, the input vector to be classified, and $f_i(\boldsymbol{X})$ the probability of $\boldsymbol{X}$ given the density function of class i. It should be pointed out that the loss due to misclassification, in most cases, cannot be observed or measured in the sample data. To avoid unnecessary complexity, it is assumed that the loss l_i is the same for all classes in our study. The PNN-guided trading strategies proposed in this study can explicitly take into account the loss (or the opportunity cost of misclassification).

Eq. (1) suggests that the Bayesian decision rule requires a knowledge of the probability density functions of possible classes. These density functions are directly estimated from a set of training samples using Parzen's window approximation method [33]. The PNN used in this study applies the

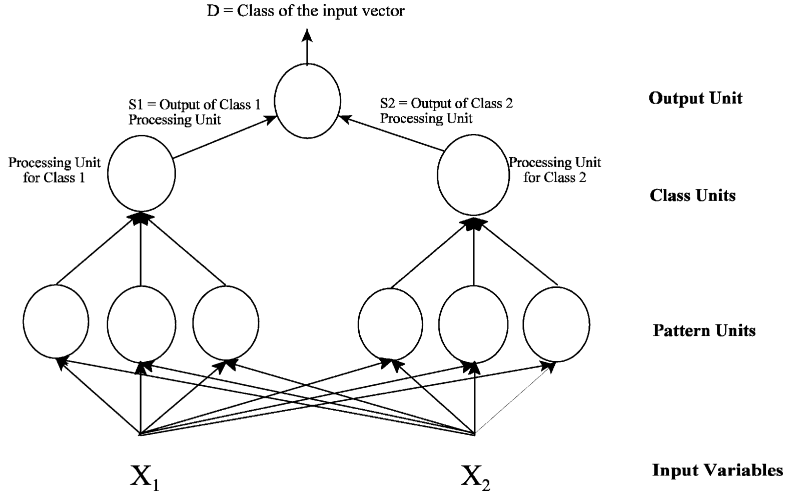

Fig. 1. An illustration of the PNN architecture.

Cacoullos's [34] multivariate extension of the original Parzen density estimation which allows us to generate the joint density functions for a set of k variables:

$$f_i(\boldsymbol{X}) = \frac{1}{(2\pi)^{k/2}\sigma^k n_i} \sum_{j=1}^{n_i} \mathrm{e}^{\frac{-(X-Y_{ij})'(X-Y_{ij})}{2\sigma^2}}, \tag{2}$$

where $\boldsymbol{X}$ is the input vector to be classified, k the number of variables in the input vector $\boldsymbol{X}, n_i$ the number of training samples which belongs to class i, $\boldsymbol{Y}_{ij}$ the jth training sample in class i, and σ is a smoothing parameter.

To implement the classification logic described above, the PNN makes use of three layers of processing units (pattern, class, and output processing units). The general construct of a typical PNN classifier is illustrated in Fig. 1. The basic PNN topology consists of four layers (an input, an output, and two hidden layers) of processing units. The input layer has a processing unit to represent each independent variable in the input vector whereas the output layer consists of a set of processing units to indicate the class affiliation. The multivariate configuration shown in Fig. 1 can be modified to a simpler univariate version by placing only one processing unit in the output layer. Each of the network models examined in this study contains a univariate output unit. The first hidden layer is called the pattern layer and uses a processing unit to "memorize" each training sample. The second hidden layer, or the class layer, is made up of an array of units with the number equal to the total number of classes. The simple PNN construct depicted in Fig. 1 represents a model with two input variables (X_1 and X_2), one univariate output (D), two classes, and three training cases for each of the two classes. Readers interested in the logic and mathematical foundation of PNN should refer to [32] for a more detailed description.

For each of the PNN models used in our study, there is a total of 68 units in the pattern layer and two units in the class layer. This configuration represents 68 cases applied to each training session and a total of two classes allowed for both directions of index returns. Since there are two classes adherent to our designated classification scheme, the processing units in the two hidden layers are

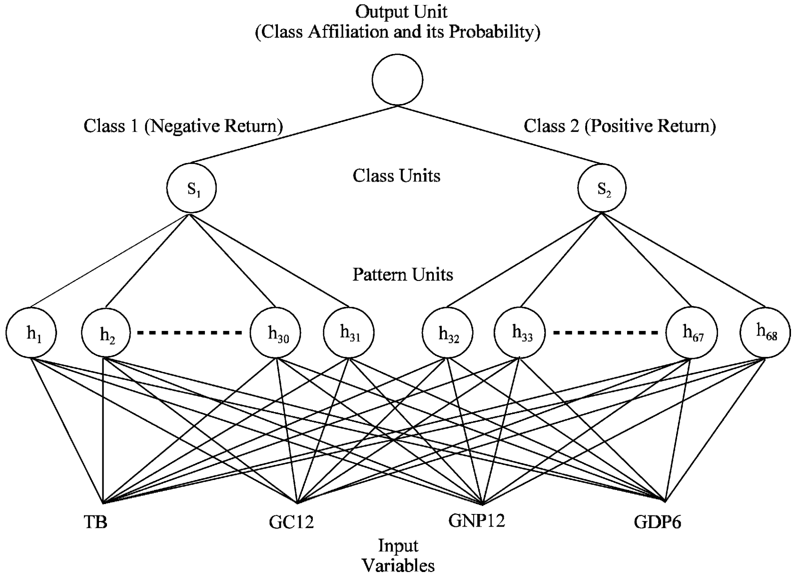

Fig. 2. PNN architecture for the first out-of-sample prediction of MR3.

divided into two subgroups, one for the negative index returns and another for the positive index returns. Fig. 2 depicts the actual PNN architecture for making the first out-of-sample forecast (Period 69) with 3-month investment horizon (MR3). Out of the total 68 samples in the training set, 31 samples have negative returns whereas 37 samples indicate positive returns. Hence, there are 31 and 37 pattern units associated with Classes 1 and 2, respectively. These pattern units are connected only to their respective class units in the subsequent hidden layer. Given the four independent variables (TB, GC12, GNP12, and GDP6), the output of this network indicates the Bayesian probability of class affiliation and thus the direction of index return. As mentioned earlier, our trading strategies will use this probability to make the dynamic asset allocation decision.

3.2.3. Training and testing scheme

A rolling horizon approach similar to that of Refenes [35] is applied to the training of our neural networks. This approach updates the training set for every out-of-sample prediction and hence incorporates the latest observed information into the network. After an out-of-sample forecast is made, the entire training set slides forward for one period and the same network training procedure is repeated. To make the first forecasts, the 68 in-sample observations (from Periods 1 to 68) are used for training. Then the trained network predicts the direction of the index return in Period 69. After the prediction is made and recorded, the training set then slides forward for one period (covering Periods 2–69) and the process of training and testing is carried out again. As a result, 60 out-of-sample forecasts on the direction of index return are generated for the purpose of testing. This training and testing scheme is applied to all 3-month (MR3), 6-month (MR6), and 12-month (MR12) models.

3.3. GMM–Kalman filter forecasting

In addition to the NN, GMM also can be used to do prediction, and explanation of GMM

Besides neural network, generalized method of moments (GMM) with Kalman filter is used to generate forecasts. There are 60 out-of-sample forecasts for each of the 3-, 6-, and 12-month investment horizons, providing an equal basis for comparison with the neural network models. The specifications for the GMM models are based on the input variables determined by the FPE minimization procedure described in Table 2. The Kalman filter estimation method is an updating method which bases the model estimates for each time period on last period's estimates plus the data for the current time period; that is, it bases its estimates only on data up to and including the current period. This makes it highly useful for constructing forecasts which are based only on historical data. When applied to a standard linear model or in our case the GMM model, it provides a convenient way to compute a new coefficient vector when an additional observation is revealed.

The Kalman filter used in this study can be written in the following way: Let $\boldsymbol{\beta}_t$ denote the vector of *states* (coefficients) corresponding to the state variables at time t. The measurement equation is the GMM model:

$$y_{t+1} = \boldsymbol{X}_i \boldsymbol{\beta}_t + \mu_t, \tag{3}$$

where y_t is the dependent variable and there are m independent variables or columns in the matrix $\boldsymbol{X}_t$. The variance of μ_t is n_t. The state vector follows the process

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + v_t \tag{4}$$

with $\mathrm{Var}(v_t) = \boldsymbol{M}_t, \mu_t$ and v_t are independent. n_t and $\boldsymbol{M}_t$ (variance of the change in state vector) are assumed to be known. The Kalman filter recursively updates the estimate of $\boldsymbol{\beta}_t$ (and its variance), using the new information in y_t and $\boldsymbol{X}_t$ for each observation. Once we have an estimate of $\boldsymbol{\beta}_{t-1}$ and its covariance matrix $\boldsymbol{\Sigma}_{t-1}$, then the updated estimate given y_t and $\boldsymbol{X}_t$ is

$$\boldsymbol{S}_t = \boldsymbol{\Sigma}_{t-1} + \boldsymbol{M}_t, \tag{5}$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{S}_t - \boldsymbol{S}_t \boldsymbol{X}_t' (\boldsymbol{X}_t \boldsymbol{S}_t \boldsymbol{X}_t' + n_t)^{-1} \boldsymbol{X}_t \boldsymbol{S}_t, \tag{6}$$

$$\boldsymbol{\beta} = \boldsymbol{\beta}_{t-1} + \boldsymbol{\Sigma}_t \boldsymbol{X}_t' n_t^{-1} (y_t - \boldsymbol{X}_t \boldsymbol{\beta}_{t-1}). \tag{7}$$

To use the previous updating equations we need to supply: $\boldsymbol{\beta}_0$, the initial state vector; $\boldsymbol{\Sigma}_0$, the initial covariance matrix of the states; n_t, the variance of the measurement equation; $\boldsymbol{M}_t$, the variance of the change in the state vector. In our study, $\boldsymbol{M}_t$, is set to $0, n_t$, $\boldsymbol{\Sigma}_0$, and $\boldsymbol{\beta}_0$ are estimated using GMM. Details of GMM estimation are described in Appendix A.

3.4. Random walk model

It is also of interest to compare the performance of the PNN and GMM models with that of the random walk models. The random walk model assumes that the best forecast is equal to the most recently observable observation. Thus, the best 3-month ahead, 6-month ahead, and 12-month ahead

Table 2
Univariate generalized method of moments regressions

	TS	TB	DS3	DS6	DS12	GC3	GC6	GC12	PC3	PC6	PC12
MR3	22.3982	−0.2394	0.0760	0.0615	0.2464	−2.6162	−4.3141	−3.7052	0.0835	0.0901	2.9410
	(0.9401)	(−3.4303)*	(0.3908)	(0.2602)	(0.7706)	(−1.3234)	(−1.1525)	(−0.6592)	(0.3270)	(0.2388)	(1.9998)*
MR6	13.5670	−0.2133	0.0206	0.0079	0.0854	−1.3434	−0.3297	−5.9956	0.0168	0.4926	2.5590
	(0.6384)	(−4.1957)*	(0.2268)	(0.0717)	(0.4717)	(−0.8453)	(−0.1232)	(−1.4479)	(0.0783)	(1.2804)	(2.3201)*
MR12	8.6385	−0.1865	0.5540	0.0793	0.1599	−0.1471	−1.8525	−8.3612	0.1383	0.3564	2.0765
	(0.8158)	(−6.3775)*	(1.3343)	(1.2018)	(1.7224)	(−0.1296)	(−0.9772)	(−2.3728)*	(0.9959)	(1.1964)	(2.4491)*

	GNP3	GNP6	GNP12	6DP3	6DP6	6DP12	CPI3	CPI6	CPI12	IP3	IP6	IP12
MR3	−0.0468	0.2000	−7.9804	−0.1323	2.2052	1.3780	−0.0578	2.9569	3.3124	−0.2002	1.4522	5.0814
	(−0.1206)	(0.3113)	(−3.0176)*	(−0.3718)	(1.9751)*	(0.4185)	(−0.0392)	(1.2190)	(1.0406)	(−0.1612)	(0.6719)	(1.4260)
MR6	0.0383	−0.4007	−6.7393	0.1192	1.5118	−1.0307	0.8345	2.6182	2.5011	0.3895	2.1666	4.0535
	(0.1125)	(−0.6447)	(−3.1790)*	(0.5627)	(1.7441)	(−0.3691)	(0.9576)	(1.4463)	(1.0482)	(0.4901)	(1.3130)	(1.5122)
MR12	−0.1472	−0.4357	−5.5598	−0.0118	−0.1948	−3.8667	0.3966	1.1288	3.4233	0.5919	1.4739	4.0326
	(−0.8106)	(−1.1750)	(−4.5487)*	(−0.0798)	(−0.3229)	(−1.6668)	(0.5054)	(0.8750)	(1.6307)	(0.9845)	(1.3565)	(1.9638)*

This table presents the results of univariate GMM regressions of various horizon excess returns of the Taiwan index on various constructed macroeconomic state variables. The data for the univariate GMM regressions cover the entire sample period. The t-statistics in parenthesis are heteroscedasticity and autocorrelation consistent. The model estimated is

$$r_{j,t} = \delta_{j,k,0} + \delta_{j,k,1}\ \mathrm{instrument}_{k,t-1} + \varepsilon_{j,t}.$$

The dependent variables are the 3-, 6-, and 12-month ahead continuously compounded annualized excess returns of the Taiwan index (MR3, MR6, and MR12). The instruments are the various independent variables described in the study and are observable on or before the last day of the month preceding the month corresponding to the first day of the forecast period.

forecast using the random walk model would be

$$\mathrm{MR3}_{t+1} = \mathrm{MR3}_t, \tag{8}$$

$$\mathrm{MR6}_{t+1} = \mathrm{MR6}_t, \tag{9}$$

$$\mathrm{MR12}_{t+1} = \mathrm{MR12}_t, \tag{10}$$

where $\mathrm{MR3}_t$, $\mathrm{MR6}_t$, and $\mathrm{MR12}_t$ are the signs (i.e., directions) of the continuously compounded 3-, 6-, and 12-month annualized excess returns of the Taiwan index previous to time t. In other words, the random walk forecasts are the sign of the most recently observable returns corresponding to our forecast horizon.

3.5. Model determination

An important aspect in developing forecasting models involves specifying the input variables. This can be done either by using: (1) economic arguments and theoretical reasoning to identify the principal determinants of the outputs; or (2) statistical testing procedure to justify the inclusion of particular input variables. In Section 2, we have given the economic arguments of why certain macroeconomic state variables may serve as inputs to our prediction models. Nevertheless, these economic arguments do not tell us how many lags of the input state variables should be included in the forecasting models. Further, establishing "pilot" models with a large number of input variables is not realistic.

Therefore, we adopt a statistical procedure using data from the in-sample estimation period (from January 1982 to August 1987) to narrow down the list of potential input state variables (see Table 1) for the prediction models and to calibrate the model specifications. The employed procedure, which is based on Akaike's minimum final prediction error (FPE), is similar to the ones used by Hsiao [36] and Kaylen [37]. The FPE criterion is derived by assuming a quadratic loss function for each equation and is computed as follows:

$$\mathrm{FPE} = \frac{\sum_{t=1}^{T} (Z_{it} - \hat{Z}_{it})^2}{T} \times \frac{T+N}{T-N}, \tag{11}$$

where Z_{it} is the ith independent state variable, N the total number of parameters in the equation, and T the number of observations. The first term on the right-hand side of Eq. (11) is a measure of modeling error while the second term, $(T+N)/(T-N)$, is an adjustment for degrees of freedom.

Assuming that the k independent state variables are considered as one group, the procedure is as follows: First, consider the dependent variable MR(d) where $d = 3, 6, 12$. Then regress MR(d) on each independent state variable, one at a time, with lags from 1 to m using GMM. The FPEs are computed by varying the variables from 1 to k with lags from 1 to m. All of the independent state variables are then searched to find the state variable, Z_i, and its associated lags, p, which lead to the minimum FPE. This procedure is repeated for each of the remaining independent variables with an additional variable or lag entering the specification only if it reduces the FPE. Table 2 tabulates the results of these estimations.

As shown in Table 2, the statistical tests suggest a model specification for MR3 which includes TB, GC12, GNP12, and GDP6 as the explanatory state variables. The MR6 model specification includes

Table 3
Input specifications of MR3, MR6, and MR12 forecasting models

Excess return for 3-month investment horizon
TB, GC12, GNP12, GDP6
Excess return for 6-month investment horizon
TB, GC12, GNP12, GDP6, CPI12
Excess return for 12-month investment horizon
TB, GC12, GNP12, GDP12, CPI12, IP12

The macroeconomic state variables for each investment horizon represent the input vector (independent variables) to GMM and PNN predictions in both in-sample estimation and out-of-sample forecasting.

TB, GC12, GNP12, GDP6, and CPI12 as the explanatory input variables. Finally, the MR12 model specification has the input vector consisting of variables TB, GC12, GNP12, GDP12, CPI12, and IP12. These model specifications are reported in Table 3.[1]

4. Results

A total of 60 out-of-sample forecasts are made for the evaluation period from September, 1987 to August, 1992. Table 4 tabulates the predicted direction of monthly index return by the PNN, the GMM–Kalman filter, and the random walk models for each of the out-of-sample periods. Actual (observed) index returns are also included for comparison.

Table 5 provides the summary statistics of the out-of-sample forecasts made by the PNN, GMM–Kalman filter, and random walk models. Shown is the table are the numbers of times (NC) the forecast correctly predicted whether the return is going to be positive or negative. It can be seen that the PNN model outperforms the others in predicting the direction of all 3-, 6-, and 12-month ahead index returns.

Further, with any length of the investment horizon, the PNN model is able to correctly predict the directions of excess returns more than 50% of the time at the 5% level of statistical significance.

[1] A quick inspection of these model specifications shows that several of the included macroeconomic input variables are likely to be collinear. However, for this study, multicollinearity in the input variables does not pose problems. First, none of the macroeconomic input variables is perfectly collinear and thus the resulting model specifications can still be estimated. Second, even in situations where multicollinearity is very high (near-multicollinearity), the OLS estimators still retain the property of BLUE. This is because near-multicollinearity per se does not violate the assumptions of the classical linear regression model. This means that unbiased, consistent estimates will be obtained, and thereby, their standard errors will be correctly estimated. The only effect of near-multicollinearity is to make it hard to get coefficient estimates with small standard errors. It is shown in Appendix A that OLS is just a special case of the GMM where the residuals are homoscedastic and do not overlap. Thus, the properties of OLS with respect to near-multicollinearity are also present when estimation is done using GMM instead of OLS. Finally, if the sole purpose of regression analysis, whether using GMM or OLS, is for forecasting rather than hypothesis testing, then near-multicollinearity is not a serious problem. Also, using a larger number of independent variables that exhibit some collinearity may at times produce forecasts that are better than a similar model that drops one or two of the collinear independent variables in an attempt to reduce multicollinearity.

Table 4
Comparison of the actual excess return with the directions predicted by each forecasting model over the out-of-sample forecast periods

Period	3-month investment horizon				6-month investment horizon				12-month investment horizon			
	Actual	*PNN*	Kalman	Rand	Actual	*PNN*	Kalman	Rand	Actual	*PNN*	Kalman	Rand
69	− 2.6193	+	+	+	− 0.6027	+	+	+	0.5837	+	+	+
70	0.1347	+	−	+	0.8023	+	-	+	0.7773	+	+	+
71	0.6393	+	+	−	0.7873	+	+	+	0.7863	+	+	+
72	1.4239	+	+	−	1.4111	+	+	+	0.7329	+	+	+
73	1.4799	+	+	+	1.5292	+	+	+	0.7224	+	+	+
74	0.9452	+	+	+	1.5274	+	+	+	0.6441	+	+	+
75	1.4084	+	+	+	1.7800	+	+	-	0.7341	+	+	+
76	1.5885	−	+	+	0.7622	+	+	+	0.5960	+	+	+
77	2.1196	+	+	+	0.7954	+	+	+	0.7369	+	+	+
78	2.1617	+	+	+	0.0646	+	+	+	0.5916	+	+	+
79	− 0.0568	+	+	+	− 0.0757	+	+	+	0.3681	+	+	+
80	− 0.5237	−	+	+	− 0.2318	−	+	+	0.2038	+	+	+
81	− 2.0274	−	+	+	− 0.3044	−	+	+	0.1394	+	+	+
82	− 0.0895	−	+	−	0.4373	+	+	+	0.4798	+	+	+
83	0.0652	−	+	−	0.6859	+	+	+	0.2695	+	+	+
84	1.4237	−	+	−	1.1260	+	+	+	0.5788	+	+	+
85	0.9690	−	+	−	0.8207	+	+	−	0.6193	+	+	+
86	1.3115	+	+	+	0.6494	+	+	−	0.4573	+	+	+
87	0.8334	+	+	+	0.5932	+	+	−	0.3228	+	+	+
88	0.6677	−	+	+	0.5226	+	+	+	0.0958	+	+	+
89	− 0.0456	−	+	+	− 0.1730	+	+	+	−0.3843	+	+	+
90	0.3107	−	−	+	− 0.0034	−	−	+	− 0.6929	−	−	+
91	0.3473	+	-	+	0.3829	−	−	+	− 0.6192	+	−	+
92	− 0.3054	+	−	−	0.2308	+	−	+	− 1.0993	−	−	+
93	− 0.3124	−	−	+	0.0199	+	-	+	− 1.4204	−	−	+
94	0.4260	−	−	+	− 0.3537	+	−	+	− 1.2565	+	−	+
95	0.7745	+	+	−	− 0.5984	−	−	−	− 0.8596	−	−	+
96	0.3622	+	−	−	− 1.3774	−	−	−	− 0.8485	−	−	+
97	− 1.1234	−	+	+	− 1.6143	−	+	+	− 1.1922	+	−	+
98	− 1.9660	−	−	+	− 2.4209	−	−	+	− 0.9352	−	−	+
99	− 3.1146	−	−	+	− 2.8507	−	−	+	− 0.8334	−	−	+
100	− 2.1027	−	−	−	− 2.1493	−	−	−	− 0.5456	+	−	+
101	− 2.8715	−	−	−	− 1.1104	−	−	−	− 0.3569	−	−	−
102	− 2.5843	−	−	−	− 0.3071	−	−	−	0.0380	−	−	−
103	− 2.1935	−	−	−	− 0.7576	−	−	−	− 0.1766	−	−	−
104	0.6555	−	−	−	0.5608	−	−	−	0.1280	−	−	−
105	1.9763	−	−	−	1.1947	−	−	−	0.4924	−	−	−
106	0.6924	+	−	−	1.0706	−	−	−	0.1848	+	−	−
107	0.4812	+	+	+	0.4090	+	+	−	− 0.0947	+	+	−
108	0.4276	+	+	+	0.3957	+	+	−	− 0.0796	−	+	−
109	1.4679	+	+	+	0.4169	+	+	−	0.1976	−	+	−
110	0.3569	+	+	+	− 0.2923	+	+	+	− 0.0736	−	+	−
111	0.3837	+	+	+	− 0.1966	+	+	+	− 0.1632	−	+	−
112	− 0.6140	+	+	+	− 0.6860	−	+	+	− 0.3703	−	+	−

Table 4 (*Continued*)

Period	3-month investment horizon				6-month investment horizon				12-month investment horizon			
	Actual	*PNN*	Kalman	Rand	Actual	*PNN*	Kalman	Rand	Actual	*PNN*	Kalman	Rand
113	− 0.9214	−	+	+	− 0.5834	−	−	+	− 0.3164	−	+	−
114	− 0.7568	−	−	+	− 0.5398	−	−	+	− 0.3380	−	+	+
115	− 0.7380	+	−	−	− 0.0056	+	−	+	− 0.3253	−	+	−
116	− 0.2255	−	+	−	0.1626	−	+	−	− 0.2334	−	+	+
117	− 0.3002	−	−	−	− 0.1099	−	−	−	− 0.4128	−	+	+
118	0.7507	−	−	−	− 0.0311	−	−	−	− 0.2761	−	+	+
119	0.5732	+	+	−	− 0.0244	−	+	−	− 0.2603	−	+	−
120	0.1026	+	+	−	− 0.1090	−	+	−	− 0.3919	−	+	−
121	− 0.7924	+	+	+	− 0.6170	−	+	−	− 0.5478	−	+	+
122	− 0.6018	−	+	+	− 0.6020	−	+	+	− 0.2369	−	+	−
123	− 0.3028	−	+	+	− 0.6908	−	+	−	− 0.0724	−	+	−
124	− 0.4257	−	+	−	− 0.4995	−	+	−	− 0.0626	−	+	−
125	− 0.5871	−	+	−	− 0.4760	−	+	−	− 0.1297	−	+	−
126	− 1.0638	−	+	−	− 0.6572	−	+	−	− 0.2017	−	+	−
127	− 0.5584	−	+	−	− 0.4661	−	+	−	− 0.1142	−	+	−
128	− 0.3499	−	+	−	0.1381	−	+	−	− 0.0914	−	+	−

Table 5
Comparison of the predictive strength of each forecasting model

	NC		
	3-month Investment horizon	6-month Investment horizon	12-month Investment horizon
PNN	44^{**}	47^{**}	50^{**}
GMM—Kalman filter	33	35^{*}	34
Random walk	32	31	38^{**}

NC denotes the number of times a forecasting model correctly predicts the direction of the index return over the 60 out-of-sample forecast periods. * and ** indicate that the number is statistically different from 30 at a 10% and 5% level of significance, respectively.

This implies that there is a less than 5% chance that these forecasting results may be due to chance alone. On the other hand, the GMM–Kalman filter model does not perform quite as well as the PNN model and is able to correctly predict the directions more than 50% of the time at the 10% level of statistical significance for only the 6-month ahead returns.

5. Trading strategies and experiment

The evidence that some of the forecasting methods documented in this study is able to correctly predict the direction of index return more than 50% of the time suggests that it may be possible to construct a set of economically profitable investment strategies. Therefore, we formulate a set of trading rules guided by the directions predicted by PNN, GMM–Kalman filter, and random walk models. The empirical testing takes the form of a trading simulation which closely mimics the timely investment decisions faced by investors in the market place. This trading simulation also allows us to evaluate the relative economic profit of the proposed investment strategies.

Essentially, the trading simulation investigates the influence of three experimental factors: (1) the length of the investment horizon; (2) the commission rates; and (3) the investment strategies. The length of investment horizon is the period of time in which the index returns are realized. This is practically the same as the horizon lengths associated with the predicted index return direction. Thus, 3-, 6-, and 12-month investment horizons are used to implement the forecasts made by the MR3, MR6, and MR12 models, respectively. Three commission rates, 0.03%, 3%, and 6%, are examined in the experiment. It should be pointed out that the 0.03% commission rate is approximately the average transaction cost faced by a typical individual investor in Taiwan. Lastly, the economic performances of the trading strategies makes use of the out-of-sample forecasts made by PNN, GMM–Kalman filter, and random walk, and a simple buy-and-hold strategy are included for comparison.

5.1. Trading simulation

We now describe the operational details of the trading simulation. The simulation experiment assumes that, at the beginning of each monthly period, the investor makes an asset allocation decision of whether to shift his liquid assets into the risk-free bonds or into the stock index fund. Liquid assets are defined as money that is currently not invested in either the risk-free bonds or the stock index fund. It should be noted that the price of the stock index fund is directly proportional to the index level. Further, it is assumed that the money that has been invested in either risk-free bonds or the stock index fund becomes liquid and will not become liquid until the end of the investor's chosen investment horizon. In other words, the invested money will become available after the selected investment horizon reaches its maturity. For example, suppose the investor has decided to use an investment horizon of 6-months. The money that he has invested into either risk-free bonds or the stock index fund in the last 6-months is considered to have been "locked up" in that security. Hence, the asset will not be available for another round of investment decision before the security matures.

The testing period runs from September 1987 to August 1992 for a total of 60 months of out-of-sample observations. The first 12 periods of the simulation test period, however, is reserved for initialization of the trading simulation. In the trading experiment, it is assumed that, during the initiation period, an investor will invest \$1 at the beginning of each month in either risk-free bonds or the stock index fund depending on his chosen investment strategy. To achieve a fair comparison of the strategies with different investment horizons, the initialization period for the 3- and 6-month investment plans are delayed so that the first maturity of any investment plan coincides with those of the 12-month investment plans which occurs at the end of Period 80. This is equivalent to the end of the 12th period in the simulation test period. The span of the initiation period varies for the different strategies in order to account for different investment horizon lengths before the first maturity. In particular, the investor will invest \$1 at the beginning of each month in either risk-free bonds or the index fund from Periods 10 through 12 if his investment strategies are based on the 3-month investment horizon. Similarly, he will invest the monthly \$1 at the beginning of each month in risk-free bonds or the index fund from Periods 7 through 12 and from Periods 1 through 12 for the

6- and 12-month investment plans, respectively. Transaction costs, brokerage cost, or commissions are assumed to be incurred every time a transaction takes place.

5.2. Trading strategies

5.2.1. PNN-guided trading strategies

Given the superior performance of the PNN forecasts, we propose two PNN-guided investment strategies which translate the predicted direction of index returns into asset allocation decision. As mentioned earlier, a distinctive characteristic of PNN is its ability to provide the Bayesian probability associated with a classification. Let P_i denote the Bayesian probability for affiliation with class i. In our proposed strategies, P_i is compared against some established threshold. Practically, if P_i is higher than the threshold level, the asset will be allocated to the security tied up to class i. This single threshold triggering can be extended to a triggering strategy involving multiple thresholds.

5.2.1.1. Single threshold triggering. There is a single threshold in this naive version of the PNN-guided trading scheme. Without any prior knowledge on the market trend or subjective preference over a certain type of security, the threshold level is set to be at 0.5. In the simulation study, the investor allocates the assets to the risk-free government bonds when the predicted return on the negative direction whereas he puts the assets into the stock index fund when the predicted return is positive. Hence, the assets will be allocated to the stock index fund if the probability of the predicted return on the up-trend is more than 0.5000. On the contrary, the assets will be allocated to the bonds if the respective probability is less than 0.5000. For the case that the probability is exactly equal to 0.5000, the asset allocation decision will follow that in the previous period.

5.2.1.2. Multiple threshold triggering. The single threshold triggering strategy ignores the asymmetric outcomes of the stock and bond markets. To be specific, the loss to misclassifying the upward movement as a downward one is not the same as the loss to misclassifying the downward movement as an upward one. If the PNN predicts an upward direction but the actual market is on the down-trend, the investor will suffer a loss in the depreciation of the stock index fund. However, if the PNN predicts a negative return in the stock market but the actual market return is positive, the investor's assets will still appreciates slightly due to the interest paid on the maturity of the risk-free bonds. To take these asymmetric payoffs into consideration, we set up a two-threshold triggering strategy. The assets will be allocated to the stock index fund if the probability of a positive predicted direction of return is at least 0.7000 while the assets will be put into the bonds if the probability is less than 0.5000. The investor will keep his asset allocation decision unchanged if the probability is between 0.5000 and 0.7000. Mathematical representations of the single and multiple threshold triggers are presented below. Eq. (12) signifies the decision rule for single threshold triggering at the 0.50 level whereas Eq. (13) represents the one for multiple threshold triggering:

$$D_t = \begin{cases} 1 & \text{for } P_2 < 0.5, \\ D_{t-1} & \text{for } P_2 = 0.5, \\ 2 & \text{for } P_2 > 0.5, \end{cases} \tag{12}$$

$$D_t = \begin{cases} 1 & \text{for } P_2 < 0.5, \\ D_{t-1} & \text{for } 0.5 \leqslant P_2 \leqslant 0.7, \\ 2 & \text{for } P_2 > 0.7, \end{cases} \tag{13}$$

where D_t is the modified result of PNN classification of the input sample for Period t, and P_2 is the PNN computed Bayesian probability that the input sample for Period t belongs to Class 2 (upward change in index level).

This investment strategy is intuitive. For a conservative investor, it would be better to invest in the risk-free bond market if the network is not certain about its prediction. In other words, the conservative investor should only invest in the riskier stock index when the network has a great confidence (i.e., a significant degree of certainty) on its prediction. The range between 0.5000 and 0.7000 represents a "buffer" region in the asset allocation decision making. There are two purposes for establishing such buffer region: (1) minimizes the number of unnecessary transactions; and (2) reduces the chance of misclassification due to uncertainty.

5.2.2. Buy and hold strategy

The investor invests his money in the stock index fund and holds the fund till the end of the simulation test horizon, that is, the end of Period 128.

5.2.3. GMM–Kalman-guided strategies

The investor will follow the directions of returns predicted by the random walk models and GMM–Kalman filter models. Similar to the learning-network-based investment strategies, the investment strategies using these econometric models allocate the assets to the stock index fund when there is a predicted up-trend and allocate the assets to the bonds when there is a predicted down-trend.

5.3. Results and analysis

The net gain in assets, number of trades executed, and the rate of return over the out-of-sample forecast horizon are shown in Table 6. Since the initial investments for various investment horizons are different, the percentage rate of return is the proper measure that can be compared across the scenarios.

$$\text{Rate of return} = \frac{\text{Net gain in assets}}{\text{Initial investment}}. \tag{14}$$

From both Table 6 and Fig. 3, it can be seen that both the single threshold (ST) and multiple threshold (MT) PNN-guided trading rules consistently outperform the ones guided by GMM-Kalman filter (KF), random walk (RW) forecasts and the buy-and-hold (BH) strategy. Also, the PNN-guided investment strategies with multiple triggering thresholds are generally better than those with single triggering threshold although the difference is marginal in some scenarios. This illustrates that the degree of certainty for PNN classification could have an impact on the interpretation of the predictions, especially when the investment horizon is long. A relatively longer investment horizon does not allow a quick switch of the underlying security and thus the strategies using multiple threshold triggering could reduce potential loss when the PNN is not certain about its prediction.

Table 6
Profits realized by various trading strategies over the out-of-sample forecast periods

	0.03% Commission					3% Commission					6% Commission				
	BH	RW	KF	ST	MT	BH	RW	KF	ST	MT	BH	RW	KF	ST	MT
3-month investment horizon: initial investment =\$1 per month for 3-months															
Net gain in assets (\$)	− 0.6731	3.9341	4.8558	10.8809	10.9077	− 0.7422	2.3061	3.6227	7.5574	7.6426	− 0.8121	1.0191	2.5482	4.9438	5.2068
No. of trades	3	28	17	27	25	3	28	17	27	25	3	28	17	27	25
% Return	− 22.44%	131.14%	161.86%	362.70%	363.59%	− 24.74%	76.87%	120.67%	251.91%	254.75%	− 27.07%	33.97%	84.94%	164.79%	173.56%
6-month investment horizon: initial investment =\$1 per month for 6-months															
Net gain in assets (\$)	− 0.4276	3.1151	9.4454	15.6901	16.5957	− 0.5931	1.9613	7.3122	13.1541	13.5415	− 0.7603	0.9233	5.4111	10.5822	10.9314
No. of trades	6	33	30	25	29	6	33	30	25	29	6	33	30	25	29
% Return	− 7.13%	51.92%	157.42%	261.50%	276.60%	− 9.89%	32.69%	121.87%	219.24%	225.69%	− 12.67%	15.39%	90.18%	180.87%	182.19%
12-month investment horizon: initial investment =\$1 per month for 12-months															
Net gain in assets (\$)	1.5988	5.2364	21.7005	24.2200	33.8743	1.1948	4.0114	18.7853	21.7753	30.6043	0.7867	2.8348	16.0163	19.4199	27.4719
No. of trades	12	31	36	26	28	12	31	36	26	28	12	31	36	26	28
% Return	13.32%	43.64%	180.84%	201.83%	282.29%	9.96%	33.43%	156.54%	181.46%	255.04%	6.56%	23.62%	133.47%	161.83%	228.93%

Legend for the investment strategies: BH, Buy and hold; RW, Random walk; KF, GMM with Kalman filter; ST, PNN with single triggering threshold at 0.5 level and MT, PNN with multiple triggering thresholds at 0.5 and 0.7 levels.

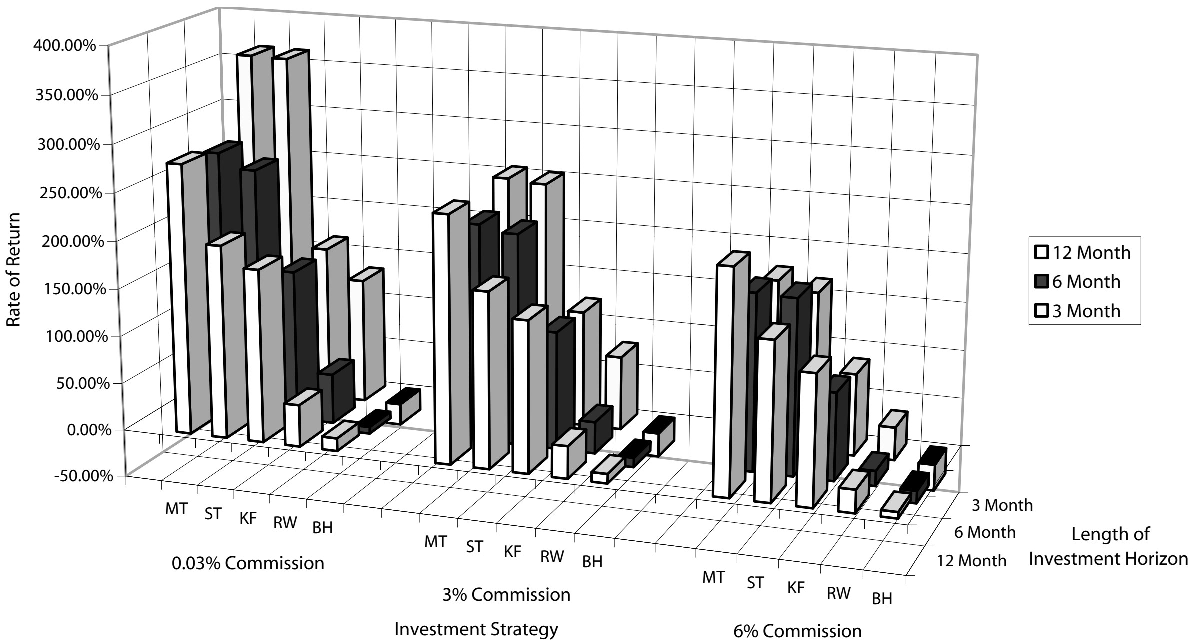

Fig. 3. Annual rate of return for each trading strategy with 3-, 6- and 12-month investment horizons.

The BH strategy results in a net loss when the investment horizon is 3- or 6-months. This is because the corresponding investment strategies, unlike the ones adopting the PNN and GMM–Kalman filter forecasts, is too rigid in that it is unable to capture the profits in the stock market and shift the realized profits into the bonds to preserve the asset worth when the market is on the down-trend.

The RW-oriented investment strategy makes the asset allocation decisions solely based on the most recent information. Thus, it can serve as a benchmark alternative in which no information prior to the previous investment period is being incorporated into the forecasts. As shown in Table 6, the performance of RW drops by at least 50% when the investment horizon stretches from 3 to 6 or 12 months. The observed sharp deterioration illustrates that this naive method which fails to take into account the possible time-series pattern is incapable of providing a reliable outlook to the more distant future. A comparison with the returns generated by PNN and GMM–Kalman filter trading strategies suggests that the timely historical information could be useful in predicting the return of the market. This notion is in agreement with many studies outlined in previous sections.

It follows the intuition that the rate of return for a given investment strategy decreases when the commission rate rises. However, it would be interesting to examine the change in the rate of return for a given trading strategy when the commission becomes more significant. From Table 6, for a particular trading strategy, the relative decrease in return is sharper for an increase in commission rate occurred in the 3-month plan than that in the 6- or 12-month plans, implying that higher commission has a greater influence on the relatively shorter investment horizon. Fig. 3 postulates that this observation is generally true for every trading strategies. A possible explanation is that the high commission harshly hampers the growth of the 3-month investment plans in early periods and thus limits the asset amount for reinvestment in later periods. However, for the plans with longer investment horizons, the assets are allowed to grow without transaction cost deduction for much longer periods of time during the early stage. This could provide a larger amount of assets for reinvestment growth at later time. In general, the investor should adopt a shorter investment horizon if the commission is low and a longer horizon if the commission is relatively higher.

The results show that the performance of both the PNN forecasts and the PNN-guided strategies are unusually strong as compared to what is normally expected from well established financial markets. The readers should be aware that Taiwan was a strong emerging economy and experienced remarkable growth during the testing period. In fact, the artificially inflated "bubble" economy created by bullish investors could have made many financial figures quite forecastable. After that, its financial sector encountered a sharp downturn and this is explained by the poor performance of the buy-and-hold strategy. In addition, the PNN's ability to estimate the underlying density function and map out the response surface may help to explain its performance in such a volatile market.

6. Conclusions

The good performance of the PNN suggests that the neural network models are useful in predicting the direction of index returns. Furthermore, PNN has demonstrated a stronger predictive power than both the GMM–Kalman filter and the random walk forecasting models. This superiority is partially attributed to PNN's ability to identify outliers and erroneous data. Compared to the other two parametric techniques examined in this study, PNN does not require any assumption of the underlying probability density functions of the class populations. Each density function is estimated by Parzen's window approximation method.

The trading experiment shows that the PNN-guided trading strategies obtain higher profits than the other investment strategies utilizing the market direction generated by the parametric forecasting methods. In addition, the PNN-guided trading with multiple triggering thresholds is generally better than the one with single triggering thresholds. The multiple threshold version is able to consider the degree of certainty of a particular PNN classification and thereby reduce potential loss in the market.

A possible extension to enhance the PNN investment decision making is to include a set of adaptive thresholds which changes dynamically in accordance with some opportunity cost. This can be achieved by setting the threshold levels with respect to the current and predicted interest rates and the price of interest rate instrument.

Acknowledgements

The authors are grateful for the helpful comments from the anonymous referees and Li D. Xu, the editor of this special issue. The authors would also like to thank Doug Blocher, Roger Schmenner, and M.A. Venkataramanan of Indiana University for their encouraging support.

Appendix A. Generalized methods of moments estimation

The econometric model used in this study is estimated using the GMM. Numerous studies, such as that by Schwert and Seguin [38], have presented evidence of heteroscedasticity in stock returns. In addition, non-normality of stock returns has been documented. The study by Blattberg and Gonedes [39] is one of the many studies which provides evidence concerning the non-normality of stock returns. Hansen [40] provides a generalized method of moments estimator (GMM) that extends White's [41] method to deal with heteroscedasticity and properly deals with serial correlation. This method was first applied by Hansen and Hodrick [42] to test the predictive power of 3-month forward rates measured at monthly intervals. Then Jorion [43] applied this method to volatility forecasts. The Hansen–White variance–covariance matrix of estimated coefficients is given by

$$\Sigma = (\boldsymbol{X}'\boldsymbol{X})^{-1}\Omega(\boldsymbol{X}'\boldsymbol{X})^{-1}, \tag{A.1}$$

where $\Omega = \boldsymbol{E}[\boldsymbol{X}' \in \varepsilon\varepsilon'\boldsymbol{X}/\boldsymbol{T}]$ is consistently estimated, using the OLS residuals ε, by

$$\hat{\Omega} = \sum_t \hat{\varepsilon}_t^2 \boldsymbol{X}_t'\boldsymbol{X}_t + \sum_s \sum_t Q(s,t)\hat{\varepsilon}_s\hat{\varepsilon}_t(\boldsymbol{X}_t'\boldsymbol{X}_s + \boldsymbol{X}_s'\boldsymbol{X}_t) \tag{A.2}$$

with $Q(s,t)$ defined as an indicator function equal to unity if there is overlap between returns at s and t, and zero otherwise. Note that in the case where the residuals are homoscedastic and do not overlap, $\boldsymbol{E}[\varepsilon_t^2] = s^2$, and $Q(s,t)$ is always zero, so that the covariance matrix collapses to the usual OLS covariance matrix, $s^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$.

References

[1] Champion SR. The great Taiwan bubble. Berkeley, CA: Pacific View Press, 1998.
[2] Banz R, Breen W. Sample-dependent results using accounting and market data: some evidence. Journal of Finance 1986;41:779–93.
[3] Jaffe J, Keim D, Westerfield R. Earnings yields, market values and stock returns. Journal of Finance 1989;44:135–48.
[4] Fama E, French K. The cross-section of expected stock returns. Journal of Finance 1992;47:427–65.
[5] Ferson WE, Harvey CR. The risk and predictability of international equity returns. Review of Financial Studies 1993;6:527–66.
[6] Jung C, Boyd R. Forecasting UK stock prices. Applied Financial Economics 1996;6:279–86.
[7] Jaffe J, Westerfield R. Patterns in Japanese common stock returns: day of the week and turn of the year effects. Journal of Financial and Quantitative Analysis 1985;20:261–72.
[8] Kato K, Ziemba W, Schwartz S. Day of the week effects in Japanese stocks. In: Elton E, Grubber M, editors. Japanese capital markets. New York: Harper & Row, 1990.
[9] Fama E, French K. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 1993;33:3–56.
[10] Fama E, Schwert W. Asset returns and inflation. Journal of Financial Economics 1977;5:115–46.
[11] Rozeff M. Dividend yields are equity risk premiums. Journal of Portfolio Management 1984;11:68–75.
[12] Keim D, Stambaugh R. Predicting returns in the stock and bond markets. Journal of Financial Economics 1986;17:357–90.
[13] Campbell J. Stock returns and the term structure. Journal of Financial Economics 1987;18:373–99.
[14] Fama E, Bliss R. The information in long-maturity forward rates. American Economic Review 1987;77:680–92.
[15] Fama E, French K. Permanent and temporary components of stock prices. Journal of Political Economy 1988;96:246–73.
[16] Fama E, French K. Dividend yields and expected stock returns. Journal of Financial Economics 1988;22:3–25.
[17] Fama E, French K. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 1990;25:23–49.
[18] Chen N, Roll R, Ross S. Economic forces and the stock market. Journal of Business 1986;59:383–403.
[19] Chan K, Chen N, Hsieh D. An exploratory investigation of the firm size effect. Journal of Financial Economics 1985;14:451–71.

[20] Harvey CR. Predictable risk and returns in emerging markets. Review of Financial Studies 1995;8:773–816.
[21] Wu Y, Zhang H. Forward premiums as unbiased predictors of future currency depreciation: a non-parametric analysis. Journal of International Money and Finance 1997;16:609–23.
[22] Aggarwal R, Demaskey A. Using derivatives in major currencies for cross-hedging currency risks in Asian emerging markets. Journal of Futures Markets 1997;17:781–96.
[23] Maberly ED. The informational content of the interday price change with respect to stock index futures. Journal of Futures Markets 1986;6:385–95.
[24] O'Connor M, Remus W, Griggs K. Going up-going down: how good are people at forecasting trends and changes in trends? Journal of Forecasting 1997;16:165–76.
[25] Hodgson A, Nicholls D. The impact of index futures markets on Australian share market volatility. Journal of Business Finance and Accounting 1991;18:267–80.
[26] Rumelhart D, McClelland J. Parallel distributed processing. Cambridge, MA: MIT Press, 1986.
[27] Caudill M, Butler C. Understanding neural networks: computer explorations. Cambridge, MA: MIT Press, 1983.
[28] Hawley D, Johnson J, Raina D. Artificial neural systems: a new tool for financial decision-making. Financial Analysts Journal 1990;46:63–72.
[29] Medsker L, Turban E, Trippi R. Neural network fundamentals for financial analysts. In: Trippi R, Turban E, editors. Neural networks in finance and investing. Chicago: Probes Publishing, 1993.
[30] Specht D. Probabilistic neural networks for classification, mapping, or associative memory. IEEE International Conference on Neural Networks, 1988;1:525–532.
[31] Specht D. Probabilistic neural networks. Neural Networks 1990;3:109–18.
[32] Wasserman P. Advanced methods in neural computing. New York: Van Nostrand Reinhold, 1993.
[33] Parzen E. On estimation of a probability density function and mode. Annals of Mathematical Statistics 1962;33:1065–76.
[34] Cacoullos T. Estimation of a multivariate density. Annals of the Institute of Statistical Mathematics 1966;18:179–89.
[35] Refenes A. Constructive learning and its application to currency exchange rate forecasting. In: Trippi R, Turban E, editors. Neural networks in finance and investing. Chicago: Probes Publishing, 1993.
[36] Hsiao C. Autoregression modeling of Canadian money and income data. Journal of the American Statistical Association 1979;74:553–60.
[37] Kaylen MS. Vector autoregression forecasting models: recent developments applied to the US hog market. American Journal of Agricultural Economics 1988;70:701–12.
[38] Schwert G, Seguin P. Heteroscedasticity in stock returns. Journal of Finance 1990;45:1129–55.
[39] Blattberg R, Gonedes N. A comparison of the stable and student distributions as statistical models of stock prices. Journal of Business 1974;47:244–80.
[40] Hansen L. Large sample properties of generalized method of moments estimators. Econometrica 1982;50:1029–54.
[41] White J. Options values under stochastic volatility: theory and empirical estimates. Econometrica 1980;48:817–38.
[42] Hansen L, Hodrick R. Forward exchange rates as optimal predictors of future spot rates. Journal of Political Economy 1980;88:829–53.
[43] Jorion P. Predicting volatility in the foreign exchange market. Journal of Finance 1995;2:507–28.

An-Sing Chen is Professor of Finance and Chairman of the Department of Finance at National Chung Cheng University, Taiwan. He received his B.B.A. in finance from Kent State University and M.B.A. in finance and Ph.D. in Business Economics from Indiana University. His areas of interest are forecasting and modeling, options and derivatives, international finance, and applied econometrics and artificial intelligence. He has published articles in a variety of journals, including *Advances in Pacific Basin Business Economics and Finance*, *Advances in Pacific Basin Financial Markets*, *Computers and Operations Research, European Journal of Operational Research, Journal of Banking and Finance, Journal of Economics and Finance*, *Journal of Investing*, *Journal of Multinational Financial Management*, *International Journal of Finance*, *International Journal of Forecasting*, *International Review of Economics and Finance*, *International Review of Financial Analysis*, and *Review of Pacific Basin Financial Markets and Policies*.

Mark T. Leung is Assistant Professor of Management Science at the University of Texas. He received his B.Sc. and M.B.A. in finance degrees from the University of California, and Master of Business and Ph.D. in Operations Management

from Indiana University. His research interests are in financial forecasting and modeling, applications of AI techniques, scheduling and optimization, and planning and control of production systems. He has published in *Decision Sciences*, *International Journal of Forecasting*, *Computers and Operations Research*, *European Journal of Operational Research*, *Journal of Banking and Finance*, *International Review of Financial Analysis*, and *Review of Pacific Basin Financial Markets and Policies*.

Hazem Daouk is an Assistant Professor of Finance, Department of Applied Economics, Cornell University. He received a DESCF from ICS Paris, France, an MBA from the University of Maryland, and a Ph.D. in Finance from Indiana University. His research interests include International Finance, Law and Finance, and Empirical methodology. He has published in the Journal of Finance, Journal of Financial Economics, International Journal of Forecasting, European Journal of Operational Research, and Computers & Operations Research.