Fac. CC. Económicas y Empresariales
Universidad de La Laguna

Fac. CC. Económicas y Empresariales
Univ. de Las Palmas de Gran Canaria

Non-linear trading rules in the New York Stock Exchange

Julián Andrada Félix *
Fernando Fernández Rodríguez*
María Dolores García Artiles *
Simón Sosvilla Rivero **

DOCUMENTO DE TRABAJO 2004-05

* Universidad de Las Palmas de G.C. Departamento de Métodos Cuantitativos en Economía y Gestión. E-mail: jandrada@dmc.ulpgc.es
** FEDEA y Universidad Complutense de Madrid.

Non-linear trading rules in the New York Stock Exchange

Julián Andrada-Félix
(Universidad de Las Palmas de Gran Canaria)
Fernando Fernández-Rodríguez
(Universidad de Las Palmas de Gran Canaria)
María Dolores García-Artiles
(Universidad de Las Palmas de Gran Canaria)
Simón Sosvilla-Rivero
(FEDEA and Universidad Complutense de Madrid)

ABSTRACT

In this paper we investigate the profitability of non-linear trading rules based on nearest neighbour (NN) predictors. Applying this investment strategy to the New York Stock Exchange, our results suggest that, taking into account transaction costs, the NN-based trading rule is superior to both a risk-adjusted buy-and-hold strategy and a linear ARIMA-based strategy in terms of returns for all of the years studied (1997-2002). Regarding other profitability measures, the NN-based trading rule yields higher Sharpe ratios than the ARIMA-based strategy for all of the years in the sample except for 2001. As for 2001, in 36 out of the 101 cases considered, the ARIMA-based strategy gives higher Sharpe ratios than those from the NN-trading rule, in 18 cases the opposite is true, and in the remaining 36 cases both strategies yield the same ratios.

JEL classification numbers: G10, G14, C53

KEY WORDS: Technical trading rules, Nearest neighbour predictors, Security markets

Introduction

In fundamental analysis, forecasts of future prices and returns are based upon economic fundamentals, such as dividends, interest, price-earning ratios, macroeconomic variables, etc.. In contrast, technical analysis looks for patterns in past prices and bases its forecasts upon extrapolation of these patterns. The basic idea is that "prices move in trends which are determined by changing attitudes of investors toward a variety of economic, monetary, political and psychological forces" (Pring, 1991, p. 2).

Although technical trading rules have been used in financial markets for over a century (see, e. g., Plummer, 1989), it is only during the last decade that technical analysis has regained the interest of the academic literature. Several authors have shown that financial prices and returns are forecastable to some extent, either from their own past or from some other publicly available information [see, e. g., Fama and French (1988), Lo and MacKinley (1988, 1997, 1999) and Pesaran and Timmerman (1995, 2000)]. Furthermore, surveys of market participants show that many use technical analysis to make decisions on buying and selling. For example, Taylor and Allen (1992) report that 90% of the respondents (among 353 chief foreign exchange dealers in London) say that they place some weight on technical analysis when forming views for one or more time horizons.

A considerable amount of work has provided support for the view that technical trading rules are capable of producing valuable economic signals in financial markets. Regarding stock markets, Brock, Hsieh and LeBaron (1992) used bootstrap simulations of various null asset pricing models and found that simple technical trading rule profits cannot be explained away by the popular statistical models of stock index returns. Later, Gençay (1996 and 1998) found evidence of non-linear predictability in stock market returns by combining simple technical trading rules and feed-forward network (see also Fernández-Rodríguez, González-Martel and Sosvilla-Rivero, 2000). As for exchange rates, Satchell and Timmermann (1995) showed that the nearest-neighbour nonlinear predictors can be implemented in a simple trading strategy which outperforms payoffs from a buy-and hold strategy based on a random walk. Later, and Fernández-Rodríguez, Sosvilla-Rivero and Andrada-Félix (2003a), considering both interest rates and transaction costs, found that a trading rule based on an NN predictor outperforms the moving average, widely used by market practicioners.

This empirical evidence has largely limited its attention to the moving average (MA) rule, which is easily expressed algebraically. Nevertheless, practitioners rely heavily on many other techniques, including a broad category of graphical methods ("heads and shoulders", "rounded tops and bottoms", "flags, pennants and wedges", etc.), which are highly non-linear and too complex to be expressed algebraically. Clyde and Osler (1997) show that the non-parametric, nearest neighbour (NN) forecasting technique can be viewed as a generalisation of these graphical methods. Based on the idea that segments of time series, taken from the past, might have a resemblance to future segments, this approach falls into a general class of models known as non-parametric regression and works by selecting geometric segments in the past of the time series similar to the last segment available before the observation we want to forecast [see Farmer and Sidorowich (1987), Härdle and Linton (1994), Cleveland and Devlin (1988) and Fernández-Rodríguez, Sosvilla-Rivero and Andrada-Félix (1997)]. Therefore, rather than extrapolating past values into the immediate future as in MA models, NN methods select relevant prior observations based on their levels and geometric trajectories, not their location in time as in the traditional Box-Jenkins (linear) methodology (see Box and Jenkins, 1976). Implicit in the NN approach is the recognition that some price movements are significant (i.e., they contribute to the formation of a specific pattern) and others are merely random fluctuations to be ignored.

Since the NN approach to forecasting is closely related to technical analysis, we aim to combine these two lines of research (non-linear forecasting and technical trading rules) to assess the economic significance of predictability in stock markets. To that end, in contrast with the previous papers, the (non-linear) predictions from NN forecasting methods are transformed into a simple trading strategy, whose profitability is evaluated against a risk-adjusted buy-and-hold strategy. Furthermore, unlike previous empirical evidence when evaluating trading performance, we will consider transaction costs, as well as a wider set of profitability indicators than those usually examined. We have applied this investment strategy to the New York Stock Exchange (NYSE), using data covering for the period January 3rd 1966 to December 31st 2002 (9312 observations).

The paper is organised as follows. Section 2 briefly presents the local NN predictors, while in Section 3 we show how the local predictions are transformed in a simple trading strategy and how we assess the economic significance of predictable patterns in the stock market. The empirical results are shown in Section 4. Finally, Section 5 provides some concluding remarks.

NN predictions

The NN method works by selecting geometric segments in the past of the time series similar to the last segment available before the observation we want to forecast [see Farmer and Sidorowich (1987) and Fernández-Rodríguez, Sosvilla-Rivero and Andrada-Félix (1997)]. This approach is philosophically very different from the Box-Jenkins methodology. In contrast to Box-Jenkins models, where extrapolation of past values into the immediate future is based on correlation among lagged observations and error terms, nearest neighbour methods select relevant prior observations based on their levels and geometric trajectories, not their location in time.

The NN forecast can be succinctly described as follows [see Fernández-Rodríguez, Sosvilla-Rivero and Andrada-Félix (1999) for a more detailed account]:

1. We first transform the scalar series x_t (t=1,...,T) into a series of m-dimensional vectors, x_i^m, t=m,...,T:

$$x_t^m = (x_t, x_{t-1}, \dots, x_{t-m+1})$$

   with *m* referred to as the *embedding dimension*. These m-dimensional vectors are often called *m-histories*.

2. Secondly, we select the *k* *m*-histories

$$x_{i_1}^m, x_{i_2}^m, x_{i_3}^m, \dots, x_{i_k}^m,$$

   that are most similar to the last available vector

$$x_T^m = \left(x_T, x_{T-1}, x_{T-2}, \dots, x_{T-m+1}\right),$$

where k=$int(\lambda T)$ (0<λ<1) with $int(\cdot)$ standing for the integer value of the argument in brackets, and where we use the subscript "i_j" (r=1,2,...,k) to denote each of the k chosen m-histories.

To that end, we search for the closest k vectors in the phase space $\Re^m$, in the sense that they maximise the function:

$$\rho\left(x_i^m, x_T^m\right)$$

(i.e., we are searching for the highest serial correlation of all m-histories, x_i^m, with the last one, x_T^m).

3. Finally, to obtain a predictor for x_{T+1}, we consider the following local regression model:

$$\hat{x}_{T+1} = \hat{\alpha}_0 x_T + \hat{\alpha}_1 x_{T-1} + \ldots + \hat{\alpha}_{m-1} x_{T-m+1} + \hat{\alpha}_m$$

whose coefficients have been fitted by a linear regression of x_{i_r+1} on $x_{i_r}^m = (x_{i_r}, x_{i_r-1}, \ldots, x_{i_r-m+1})$ $(r = 1, \ldots k)$. Therefore, the $\hat{\alpha}_i$ are the values of α_i that minimise

$$\sum_{r=1}^{k} (x_{i_r+1} - \alpha_0 x_{i_r} - \alpha_1 x_{i_r-1} - \ldots - \alpha_{m-1} x_{i_r-m+1} - \alpha_m)^2$$

Note that the NN predictors depend on the values of the embedding dimension m and the number of closest k points in the phase space $\Re^m$. Although there are some heuristic methods that have been proposed in the literature to choose these key parameters (see Fernández-Rodríguez, Sosvilla-Rivero and Andrada-Félix, 2003b), we make use of genetic algorithms (GA) to jointly determine the optimal values for m and k.

GA, developed by Holland (1975), are a class of adaptive search and optimisation techniques that have the advantage of being able to evaluate loss functions associated with the predictor parameters with no assumption regarding the continuity or differentiability of the loss function.

Furthermore, the use of GA allows us to mitigate the danger of a “data snooping” bias (i.e., the substantial danger of detecting spurious patterns in security returns if trading strategies are both discovered and tested in the same database). We use the cross-validation method when choosing the key parameters in the NN predictors. This method, widely used in non-parametric regression (see, e.g., Efrom, 1983), consists in allowing the data to select the key parameters. To that end, the sample is divided in two sub-samples. Sub-sample I is usually called the “training period” and is used to choose the parameters that minimise some loss function defined in terms of prediction errors. Finally, the model evaluation is performed, using the sub-sample II, known as the “validation period”. The performance of the model can only be judged in the validation period, not in the training period.

To that end, we select a validation period $V = \{x_t : N < t \leq T\}$ for some $N < T$. For each $x_t \in V$ we obtain a one-step ahead prediction $\hat{x}_{t+1}$ for the observation x_{t+1} using only past information: $\hat{x}_{t+1} = E(x_t / F_t)$, where the past information set is $F_t = \{x_s : 1 \leq s \leq t\}$. This allows us to select the key parameters using some measure of forecasting accuracy, such as the root mean square prediction error:

$$R_k(m) = \sqrt{\sum_{x_t \in V} \left(\hat{x}_t - x_t\right)^2}$$

Therefore, using a GA we select that pair (m,k) that minimises $R_k(m)$. This approach is similar to that employed by Casdagli (1992a and b), who proposed a procedure based on the behaviour of the mean square prediction error (normalised by the standard deviation of the time series being predicted) associated with NN predictors with different values for m and k. Nevertheless, Casdagli used this algorithm in a different setting, since his objective wasto distinguish between low-dimension chaotic behaviour and stochastic linear behaviour by comparing short-term predictions.

A GA is initiated with a population of randomly generated solution candidates, which are evaluated in terms of an objective function. These candidates are usually represented by vectors consisting of binary digits. Promising candidates, as represented by relatively better performing solutions, are then combined through a process of binary recombination referred to as crossover. Finally, random mutations are introduced to safeguard against the loss of genetic diversity, avoiding local optima. Successive generations are created in the same manner and evaluated using the objective function until a well-defined criterion is satisfied.

In order to determine which solution candidates are allowed to participate in the crossover and then to undergo possible mutation, we apply the genitor selection method proposed by Whitley (1989). This approach involves ranking all individuals according to performance and then replacing the poorly performing individuals by copies of better performing ones. In addition, we apply the commonly used single point crossover, consisting of randomly pairing candidates surviving the selection process and randomly selecting a break point at a particular position in the binary representation of each candidate. This break point is used to separate each vector into two subvectors. The two subvectors to the right of the break point are exchanged between the two vectors, yielding two new candidates. Finally, mutation occurs by randomly selecting a particular element in a particular vector. If the element in question is a one it is mutated to zero, and viceversa. This occurs with a very low probability in order not to destroy promising areas of search space.

3. Trading rules

The trading rule considered in this study is based on a simple market timing strategy, consisting of investing total funds in either the stock market or a risk free security. The forecast from NN predictors is used to classify each trading day into periods "in" (earning the market return) or "out" of the market (earning the risk-free rate of return). The trading strategy specifies the position to be taken the following day, given the current position and the "buy" or "sell" signals generated by the NN. On the one hand, if the current state is "in" (i. e., the investor is holding shares in the market) and the share prices are expected to fall on the basis of a sell signal generated by the NN predictor, then shares are sold and the proceeds from the sale are invested in the risk free security [earning the risk-free rate of return $r_f(t)$]. On the other hand, if the current state is "out" and the NN predictor indicates that share market prices will increase in the near future, the rule returns a "buy" signal and, as a result, the risk free security is sold and shares are bought [earning the market rate of return $r_m(t)$]. Finally, in the other two cases, the current state is preserved.

The trading rule return over the entire period of 1 to T can be calculated as:

$$r = \sum_{t=1}^{T} r_m(t) \;\; I_b(t) + \sum_{t=1}^{T} r_f(t) \;\; I_s(t) + n \;\; \log \frac{1-c}{1+c}$$

where $r_m(t) = \log P_t - \log P_{t-1}$ is the market rate of return, P_t is the closing price (or level of the composite stock index) on day *t;* $I_b(t)$ and $I_s(t)$ are indicator variables equal to one when the NN predictor signal is, respectively, “buy” and “sell”, and zero otherwise, satisfying the relation $I_b(t) x I_s(t) = 0$, $\forall t \in [1, T]$; n is the number of transactions; and c denotes the one-way transaction costs (expressed as a fraction of the price).

In order to assess profitability, it is necessary to compare the return from the trading rule based on the NN predictor to an appropriate benchmark. To that end, we construct a weighted average of the return from being long in the market, and the return from holding no position in the market and thus earning the risk free rate of return (Allen and Karjalainen, 1999). The return on this risk-adjusted buy-and-hold strategy can be written as

$$r_{bh} = \beta \sum_{t=1}^{T} r_m(t) + (1-\beta) \sum_{t=1}^{T} r_f(t) + 2 \log \frac{1-c}{1+c}$$

where β is the proportion of trading days that the rule is in the market.
As a further profitability assessment, we also consider a linear ARIMA(1,1,0) predictor and use it to generate “buy” or “sell” signals in the same way we have described for the NN predictor, computing its excess return from a risk-adjusted buy-and-hold strategy.

In the empirical implementation, we will modify the simple rule introducing a filter in order to reduce the number of false buy and sell signals by eliminating “whiplash” signals when the NN or the ARIMA predictor at date *t* is close to the closing price at *t-1*. The filter can be interpreted as a representing the risk that the investor is willing to assume. The filtered rule will generate a buy (sell) signal at date *t* if the NN or the ARIMA predictor is greater than (is less than) the closing price at *t-1* by a percentage δ

of the standard deviation σ of the first difference of the price time series from *1* to *t-1*. Therefore, if $\hat{P}_t$ denotes the NN or the ARIMA prediction for P_t:

- If $\hat{P}_t > P_{t-1} + \delta\sigma$ and we are out the market, a buy signal is generated. If we are in the market, the trading rule suggests that we continue in the market.
- If $\hat{P}_t \leq P_{t-1} - \delta\sigma$ and we are in the market, a sell signal is generated. If we are out of the market, we continue holding the risk free security.

4. Data and preliminary results

The data consists of the daily closing values of the NYSE Composite Index, which reflects the price of all common stocks listed on the New York Stock Exchange. The data is collected over the period January 3rd 1966 to December 31st 2002, consisting of 9312 observations (see Figure 1)[1].

[Insert Figure 1, here]

Table 1 provides summary statistics of the price levels and returns series. As can be seen, the series are positively skewed and strongly serially correlated. The Jarque-Bera (1980) test for joint normal kurtosis and skewness rejects the normality hypothesis and the Box-Pierce *Q*-statistic indicates significant autocorrelation. Regarding the augmented Dicky-Fuller test, while we are unable to reject a unit root for the price level, we do reject it for the returns series.

[Table 1 here]

Before computing our NN predictors, we have tested for the presence of nonlinear dependence in the series, since evidence of nonlinearity would support our approach to forecasting. To that end, we used a simpler test procedure by calculating the BDS test statistic (Brock, Dechert and Scheinkman, 1987). It is based on the concept of correlation integral:

---

[1] The data are taken from http://www.nyse.com/marketinfo/nysestatistics.html#Indices

$$C^m(\varepsilon) = \frac{2}{T_N(T_N - 1)} \sum_{i<j} I_\varepsilon(x_i^m, x_j^m)$$

where T_N=T-m+1 is the number of m-histories that can be made from a sample size T and $I_\varepsilon(x_i^m, x_j^m)$ is an indicator function that equals one if $\left\| x_i^m - x_j^m \right\| < \varepsilon$ and zero otherwise, where $\|\cdot\|$ is the L_∞ norm on $\Re^{\mathrm{m}}$. Therefore, the correlation integral is an estimate of the probability that any two m-histories (x_i^m and x_j^m) in the series are near to each other, where nearness is measured in terms of the distance between them being less than ε. Under the null hypothesis that x_t is independent and identically distributed (iid):

$$C^m(\varepsilon) \to C^1(\varepsilon)^m \text{ as } T \to \infty$$

from which the BDS test statistic is defined as:

$$BDS(m, \varepsilon) = \frac{\sqrt{T_N}\left[C^m(\varepsilon) - C^1(\varepsilon)^m\right]}{\hat{\sigma}_m(\varepsilon)}$$

where $\hat{\sigma}_m(\varepsilon)$ is an estimate of the standard deviation of $C^m(\varepsilon) - C^1(\varepsilon)^m$. Brock, Dechert and Scheinkman (1987) show that, under the null hypothesis of iid, the *BDS(m, ε)* statistic converges in distribution to a standard normal variable with unit variance [i.e., N(0,1)] as $T \to \infty$.

Note that the BDS test statistic depends on the value of the embedding dimension and the distance between the standard deviation of the data (m and ε, respectively). Following Hsieh (1989) and Brock, Hsieh and LeBaron (1991), we use values of m from 2 to 8 (i.e., we form vectors using observations of 2 to 8 consecutive trading days), and values for ε ranging from 0.75σ to 2.75σ, where σ denotes the standard deviation of the series. This range of values for ε was chosen in an attempt to avoid the situation where ε is too small and no m-histories are “close” and, conversely, to avoid the situation where ε is too big and all m-histories are “close” to one another. As for the practical implementation of the test, here this is done by using the residuals of an AR(p) model to remove linear dependence in the times series[2]. The results are reported in Table 2 (upper part)[3].

[Table 2 here]

With over 9000 observations, we can use the normal standard tables to assess significance, since the small sample properties only become important for sizes less than 500 (see Brock, Hsieh and LeBaron, 1991). As can be seen in Table 2, the null hypothesis of *iid* is always rejected at the one percent marginal significance level. These results are in line with those of Ramsey (1990), Hsieh (1991) and Fernández-Rodríguez, Sosvilla- Rivero and García-Artiles (1997, 1999), among others. This opens alternatives of nonlinear dependence and also nonstationarity. Although nonstationarity of the series is detected, in this paper we explore the use of nonlinear dependencies in order to forecast the series.

In order to reinforce our results, we follow Scheinkman and LeBaron's (1989) suggestion, recreating the data series by sampling randomly without replacement from the data until one has a "shuffled" series of the same length as the original. The shuffled series should be completely random (though preserving the original distribution). Applying the BDS test to the shuffled residuals series, the null hypothesis of *iid* is retained, because all BDS test values are less than the critical values (see lower part of Table 2). Therefore, there is evidence that some of the nonlinear structure present in the original series has been removed by shuffling.

5. Assessing the profitability of non-linear trading rules based on NN predictors

[2] Using Akaike (1974) information criterion, the number of lags chosen to select the most parsimonious AR(p) model was 6.

[3] We only report the results for the period running from 3 January 1966 to 31 December 2001, since the results for the other subperiods were qualitatively similar. The results for the subperiods 3 January 1966 to 31 December 1996, 3 January 1966 to 31 December 1997, 3 January 1966 to 31 December 1998, 3 January 1966 to 31 December 1999, 3 January 1966 to 31 December 2000, and 3 January 1966 to 31 December 2001 are available from the authors upon request.

Based on the indications of nonlinearities previously reported, we proceeded to assess the performance of trading rules using our NN predictors for the NYSE Composite Index. In order to avoid the possibility of data-snooping in the selection of the sample period analysed, we consider the performance of our NN predictors in five successive years (1997-2001). In each case, we use data on the previous year to the one we are going to forecast to select by GA the parameters *m* and *k* that minimise the root mean square prediction error. Once these parameters are selected, we look for the *k* *m*-histories most similar to the last available vector using all past histories of the time series. Therefore, in order to forecast the year 1997, we use the in-sample training period running from January 2nd 1996 to December 31st 1996, while the out-of-sample validation period covers the period from January 2nd 1997 to December 31st 1997.

As can be seen in Figure 2, although the NYSE Composite Index rose early in the year 1997, it fell with the upturn in interest rates in February. As interest rates subsequently declined and earnings reports remained quite upbeat, the market again advanced, reaching new highs in the spring. However, the advances were much more modest, on balance, over the second half of the year, particularly after October when the increasing difficulties in Asia led investors to lower expectations for the earnings of some U. S. firms. As for 1998, the NYSE Composite Index began at 511.19 and reached 39 new all-time highs before closing the year at a 595.81 (an increase of 16.6%) (see Figure 3). The index experienced a high degree of volatility: during the first part of the year significant increases were observed, followed by a outbreak of turbulence in the financial markets of emerging economies that wiped out almost all of the market gains through mid-July. However, global financial markets changed once again during the last quarter of 1998, with the stock prices showing a significant recovery after the first week of October 1998. Regarding 1999, the NYSE Composite Index had a gain of 9.2% to close at 650.30, after reaching 20 new all-time highs (see Figure 4). Optimism about long-term earnings growth prospect for high-technology firms played an important role in this increase in stock prices. As is shown in Figure 5, the NYSE Composite Index reached its record high of 677.58 on 1 September 2000 before closing the year at 656.87 (an increase of 1.01%). The implosion of Internet-related companies, high oil prices and a tighter Fed policy and uncertain political environment stemming from the prolonged Presidential election contributed to the discontinuance of the upward trend that had been apparent from 1995 onwards. Finally, despite the monetary easing throughout 2001, stock prices were negatively affected by the further slowdown in the pace of economic activity in the course of the year. Fears of an economic recession turned to reality by the first quarter of 2001, which curtailed profitability for most companies and dampened stock prices. This trend was only exacerbated in the wake of the September 11th terrorist attacks. Although the market gained some momentum in the fourth quarter as the monetary and fiscal measures adopted in the United States helped to restore market confidence, it was not enough to correct a negative performance for the year, and the NYSE Composite Index lost 10.2% to close at 589.80 by year-end (see Figure 6). At the beginning of 2002, the NYSE showed a notable change in performance, mainly due to the favourable analysis presented by the chairman of the US Reserve Board on the US economy at a Senate hearing. After some period of high volatility, with ups and downs in prices taking place without any let-up and leading to high risk investments, there was a sharp drop in mid-May due to the uncertainty about US recovery and corporate profits and accounting practices, followed by a recovery in August spurred by the agreement reached between the US Congress and the Senate to pass a law to deal with corporate fraud. Finally, after the initial lows reached in the second week of October, the market moved clearly upward in view of expectations created by the imminent drop in the cost of money and some relaxation of the threat of war with Iraq. Nevertheless, once interest rates had been changed, the publication at the beginning of December of figures about the US economy raised new feelings of uncertainty in the market. As a result, the market tended to backtrack from the increases in the previous two months which, while not especially painful, prevented any consolidation of the recovery of the market from annual lows and thus, at least partially, eased the poor results of a year that was definitely negative, with the NYSE Composite Index closing at 472.54 (a lost of 10,2%) (Figure 7). As can be seen, the forecasting period is very challenging, since in our sample there are alternating episodes of generally rising or generally falling prices (so-called “bull” and “bear” markets), not only from one period to the other, but also within each year to be forecast.

[Insert Figures 2 to 7, here]

In Table 3 we report the values of *m* and *k* selected by GA. It is interesting to note that the embedding dimension is relatively low (between 1 and 5), except for 1998 when *m*=10, reflecting perhaps the Asian crisis in 1997 in the sense of increasing the amount of information needed for capturing the underlying dynamics of the time series. Regarding the number of closed *k* points, it is always chosen between 1% and 8% of the sample.

[Insert Table 3, here]

We have computed the filter technique for 101 filters ranging in size from 0% to 1%. Regarding the transaction costs, results by Sweeny (1988) suggest that in the mid-1970s large institutional investors could achieve one-way transaction costs in the range of 0.1-0.2%. Even thought there have been substantial reductions in costs in the last decades, we use one-way transaction costs of 0.15%. As for the risk-free rate of return, following the literature we use the three-month Treasury-Bill Rate (see, e. g., Bodie, Kane and Marcus, 2002)[4].

The out–of-sample net return statistics are plotted in Figures 8 to 13. For 1997, the trading strategy based on the NN predictor always renders positive net returns, being higher than those obtained from the ARIMA trading rule in 88 out of the 101 possible cases[5], than those from a simple buy and hold strategy in 5 out of the 101 cases[6], and than a risk-adjusted buy-and-hold strategy in 72 out of the 101 cases[7]. For 1998, once again the NN trading rule always produces positive net returns, being higher than those obtained from the ARIMA trading rule in 77 out of the 101 possible cases[8], than those from a simple buy and hold strategy in 69 out of the 101 cases[9], and than a risk-adjusted

[4] The data are taken from http://research.stlouisfed.org/fred/data/wkly/dtb3

[5] We find higher net returns from the trading rule based on the NN predictors than those from an ARIMA-based strategy when the filter takes the values 0.00-0.04, 0.07-0.12, 0.22-0.28 and 0.31-1.00.

[6] We find higher net returns from the trading rule based on the NN predictors than those from a simple buy-and-hold strategy when the filter takes the values 0.24 and 0.36-0.39. For 1997, the net return from the simple buy-and-hold strategy is 0.27.

[7] We find higher net returns from the trading rule based on the NN predictors than those from a risk-adjusted buy-and-hold strategy when the filter takes the values 0.01, 0.21-0.24, and 0.33-1.00.

[8] We find higher net returns from the trading rule based on the NN predictors than those from an ARIMA-based strategy when the filter takes the values 0.12-0.13 and 0.26-1.00.

[9] We find higher net returns from the trading rule based on the NN predictors than those from a simple buy-and-hold strategy when the filter takes the values 0.23-0.71 and 0.73-0.92. For 1998, the net return from the simple buy-and-hold strategy is 0.15.

buy-and-hold strategy in 81 out of the 101 cases[10]. For 1999, the NN trading rule produces positive net returns for all positive values of the filter, being higher than those obtained from the ARIMA trading rule in 88 out of the 101 possible cases[11], than those from a simple buy and hold strategy in 77 out of the 101 cases[12], and than a risk-adjusted buy-and-hold strategy in 81 out of the 101 cases[13]. For 2000, the NN trading rule produces positive net returns in 85 out of the 101 possible cases, being higher than those obtained from the ARIMA trading rule in 48 out of the 101 possible cases[14], than those from a simple buy and hold strategy in 63 out of the 101 cases[15], and than a risk-adjusted buy-and-hold strategy in 45 out of the 101 cases[16]. For 2001, the NN trading rule produces positive net returns in 62 out of the 101 possible cases, being higher (equal) than those obtained from the ARIMA trading rule in 18 (46) out of the 101 possible cases, higher than those from a simple buy and hold strategy in 87 out of the 101 cases[17], and higher (equal) than a risk-adjusted buy-and-hold strategy in 36 (47) out of the 101 cases[18]. Finally, for 2002 the NN trading rule produces positive net returns in 64 out of the 101 possible cases, being higher than those obtained from the ARIMA trading rule in 77 out of the 101 possible cases[19], than those from a simple buy and hold

---

[10] We find higher net returns from the trading rule based on the NN predictors than those from a risk-adjusted buy-and-hold strategy when the filter takes the values 0.13 and 0.21-1.00.

[11] We find higher net returns from the trading rule based on the NN predictors than those from an ARIMA-based strategy when the filter takes the values 0.00-0.01, 0.09-0.13 and 0.20-1.00.

[12] We find higher net returns from the trading rule based on the NN predictors than those from a simple buy-and-hold strategy when the filter takes the values 0.24-1.00. For 1999, the net return from the simple buy-and-hold strategy is 0.09.

[13] We find higher net returns from the trading rule based on the NN predictors than those from a risk-adjusted buy-and-hold strategy when the filter takes the values 0.13, 0.18 and 0.22-1.00.

[14] We find higher net returns from the trading rule based on the NN predictors than those from an ARIMA-based strategy when the filter takes the values 0.08-0.55.

[15] We find higher net returns from the trading rule based on the NN predictors than those from a simple buy-and-hold strategy when the filter takes the values 0.09-0.12, 0.14-0.58 and 0.87-1.00. For 2000, the net return from the simple buy-and-hold strategy is 0.02.

[16] We find higher net returns from the trading rule based on the NN predictors than those from a risk-adjusted buy-and-hold strategy when the filter takes the values 0.11, 0.14 and 0.16-0.58.

[17] We find higher net returns from the trading rule based on the NN predictors than those from a simple buy-and-hold strategy when the filter takes the values 0.10-0.22 and 0.27-1.00. For 2001, the net return from the simple buy-and-hold strategy is -0.09.

[18] We find higher net returns from the trading rule based on the NN predictors than those from a risk-adjusted buy-and-hold strategy when the filter takes the values 0.12-0.17, 0.20-0.22 and 0.27-0.58.

[19] We find higher net returns from the trading rule based on the NN predictors than those from an ARIMA-based strategy when the filter takes the values 0.00-0.12, 0.15-0.19, 0.25, 0.38-0.45 and 0.51-1.00.

strategy in all cases[20], and than a risk-adjusted buy-and-hold strategy in 96 out of the 101 cases[21].

[Insert Figures 8 to 13, here]

According to the results in Figures 8 to 13, it appears that a strategy of using a filter between 0.27 and 0.53 could be recommended for practitioners when applying this non-linear trading rule.

Given that individuals are generally risk averse, besides the excess return, we also consider the Sharpe ratio (Sharpe, 1966). This is a risk-adjusted return measure given by:

$$RS = \frac{\overline{r}}{\sigma}$$

where $\overline{r}$ is the average annualised return of the trading strategy and σ is the standard deviation of daily trading rule returns. As can be seen, higher Sharpe ratios correspond to higher mean annual net returns and lower volatility.

Figures 14 to 19 show the Sharpe ratios for the trading rule based on the NN and ARIMA predictors. For all out-of-sample periods, the NN-based trading rule yields higher Sharpe ratios than the ARIMA-based strategy in more than 80 out of the 101 cases considered, except for 1998 and 2001 (74 out of 101 cases and 18 out of 101 cases, respectively)[22]. Furthermore, it should be observed that the non-linear trading rule always generates positive Sharpe ratios in 1997, 1998 and 1999.

---

[20] For 2002, the net return from the simple buy-and-hold strategy is -0.23.

[21] We find higher net returns from the trading rule based on the NN predictors than those from a risk-adjusted buy-and-hold strategy when the filter takes the values 0.00-0.02. 0-07-0.26 and 0.28-1.00.

[22] For 1997, the NN-based trading rule yields higher Sharpe ratios than the ARIMA-based strategy in 87 out of the 101 cases considered, while the opposite is true in the 14 remained cases. For 1998, the NN-based trading rule yields higher Sharpe ratios than the ARIMA-based strategy in 74 out of the 101 cases considered, while the opposite is true in the 27 remained cases. For 1999 , the NN-based trading rule yields higher Sharpe ratios than the ARIMA-based strategy in 88 out of the 101 cases considered, while the opposite is true in the 13 remained cases. For 2000, the NN-based trading rule yields higher Sharpe ratios than the ARIMA-based strategy in 84 out of the 101 cases considered, while the opposite is true in the 17 remained cases. For 2001, the NN-based trading rule yields higher Sharpe ratios than the ARIMA-based strategy in 18 out of the 101 cases considered, the ARIMA-based trading rule yields higher Sharpe ratios than the NN-based strategy in 36 out of the 101 cases considered, and in the 47 remained cases, both strategies yield the same Sharpe ratio. Finally, for 2002, the NN-based trading rule yields higher

[Insert Figures 14 to 19, here]

Sharpe ratios than the ARIMA-based strategy in 82 out of the 101 cases considered, while the opposite is true in the 19 remained cases.

6. Concluding remarks

The purpose of our paper has been to contribute to the debate on the relevance of non-linear forecasts of high-frequency data in financial markets. To that end, we have presented the results of applying the nearest neighbour (NN) predictors introduced by Farmer and Sidorowich (1987) and Fernández-Rodríguez, Sosvilla-Rivero and Andrada-Félix (1997) to the New York Stock Exchange (NYSE), using data for the period January 3rd 1966 to December 31st 2002. The NN predictors have been transformed into a simple trading strategy, whose profitability is evaluated against a risk-adjusted buy-and-hold strategy. In doing so, our approach incorporates the essence of technical analysis: to identify approach regularities in the time series of prices by extracting non-linear patterns from noisy data. Furthermore, unlike previous empirical evidence, when evaluating trading performance, we have not only considered transaction costs and therefore net returns, but also Sharpe ratios as additional profitability indicators.

The main results are as follows. Even though the forecasting period is very heterogeneous, with alternating episodes of "bull" and "bear" markets (not only from one period to the other, but also within each year to be forecast), the NN-based trading rule produces positive net returns in 513 out of the 606 cases considered. The NN-based trading rule is superior in terms of net returns to both a risk-adjusted buy-and-hold strategy and an ARIMA-based strategy for all of the years studied, except for 2000 and 2001. Nevertheless, for 2001 in 47 (46) out of the 101 cases the NN-based strategy produces equal net return than those from a risk-adjusted buy-and-hold (ARIMA-based) strategy. In addition, the NN-based trading rule produces higher net returns than those from a simple buy-and-hold strategy, except for 1997. Regarding other profitability measures, the NN-based trading rule yields higher Sharpe ratios than the ARIMA-based strategy for all of the years in the sample except for 2001, when in 36 out of the 101 cases considered, the ARIMA-based strategy gives higher Sharpe ratios than those from the NN-trading rule, in 18 cases the opposite is true, and in the remaining 36 cases both strategies yield the same ratios.

The results in this paper indicate that there exists potential for investors to generate excess returns in stock markets by adopting technical trading rules based on NN predictors.

References:

Akaike, H. (1974): “A new look at the statistical model identification”, *IEEE Transactions of Automatic Control*, 19, 716-723.

Allen, F. and Karjalainen, R. (1999): "Using genetic algorithms to find technical trading rules", *Journal of Dinancial Economics*, 51, 245-271.

Bodie, Z. , Kane, A. and Marcus, A. J. (2002): *Essentials of Investments.* New York: MacGraw-Hill, 5th edition.

Box, G. E. P. and G. M. Jenkins (1976): *Time Series Analysis Forecasting and Control*. San Francisco: Holden-Day.

Brock, W. Hsieh, D. A. and B. LeBaron (1991): *Nonlinear Dynamics, Chaos and Instability.* Cambridge, MA: MIT Press.

Brock, W., Dechert, W. and J. Scheinkman (1987): “A test for independence based on the correlation dimension”. University of Wisconsin, Economics Working Paper SSRI-8702.

Brock, W., Lakonish, J. and B. LeBaron (1992): “Simple technical rules and the stochastic properties of stock returns”, *Journal of Finance*, 47, 1731-1764.

Cleveland, W. S. and Devlin, S. J. (1988): "Locally weigthed regression: An approach to regression analysis by local fitting", *Journal of the American Statistical Association* 83, 596-610.

Clyde, W. C. and C. L. Osler (1997): "Charting: Chaos theory in disguise?", *Journal of Future Markets*, 17, 489-514.

Fama, E. and K. French (1988): “Dividends yield and expected stock returns*”, Journal of Financial Economics*, 22, 3-27.

Farmer, D. and J. Sidorowich (1987): “Predicting chaotic time series”, *Physical Review Letters*, 59, 845-848.

Fernández-Rodríguez, F., González-Martel, C. and S. Sosvilla-Rivero (2000): “On the profitability of technical trading rules based on artifitial neural networks: Evidence from the Madrid stock market”, *Economics Letters*, 61, 89-94.

Fernández-Rodríguez, F., Sosvilla-Rivero, S. and M. D. García-Artiles (1997): "Using nearest neighbour predictors to forecast the Spanish stock market*", Investigaciones Económicas*, 21, 75-91.

Fernández-Rodríguez, F., Sosvilla-Rivero, S. and M. D. García-Artiles (1999): "Dancing with bulls and bears: Nearest-neighbour forecasts for the Nikkei Index", *Japan and the World Economy*, 11, 395-413.

Fernández-Rodríguez, F., Sosvilla-Rivero, S. and J. Andrada-Félix (1999): “Exchange-rate forecasts with simultaneous nearest-neighbour methods: Evidence from the EMS”, *International Journal of Forecasting*, 15, 383-392.

Fernández-Rodríguez, F., Sosvilla-Rivero, S. and J. Andrada-Félix (2003): “Technical Analysis in Foreign Exchange Markets: Evidence from the EMS”, *Applied Financial Economics*, 13, 113-122.

Fernández-Rodríguez, F., Sosvilla-Rivero, S. and J. Andrada-Félix (2003b): "Nearest-neighbour predictions in foreign exchange markets", in Shu-Heng Chen and Paul Wang, eds., *Computational Intelligence in Economics and Finance*, Berlin: Physica Verlag.

Gençay, R. (1996) "Nonlinear prediction of security returns with moving average rules", *Journal of Forecasting*, 15, 165-174.

Gençay, R. (1998) "The predictability of securities returns with simple technical rules", *Journal of Empirical Finance*, 5, 347-359.

Härdle, W. and Linton, O. (1994): "Applied nonparametric methods", R. F. Engle. and D. Mac Fadden, eds., in *Handbook of Econometrics*, Amsterdam: Elsevier, Vol. 4.

Hsieh, D. A. (1989): “Testing for nonlinear dependence in daily foreign exchange rates”.
Journal of Business, 62, 339-368.

Hsieh, D. A., (1991): “Chaos and nonlinear dynamics: Application to financial markets”. *Journal of Finance*, 46, 1839-1877.

Lo, A. W. and A. C. MacKinley (1988): “Stock market prices do not follow random walks: Evidence from a simple specification test”, *Review of Financial Studies*, 1, 41-66.

Lo, A. W. and A. C. MacKinley (1997): “Maximizing predicibility in stock and bonds markets”, *Macroeconomic Dynamics*, 1, 102-134.

Lo, A. W. and A. C. MacKinley (1999): *A non-random walk down Wall Street.* Princeton: Princeton University Press.

MacKinnon, J. G. (1991): “Critical values for cointegration tests”, in Robert F. Engle and Clive W. J. Granger, eds., *Long-run Economic Relationships: Readings in Cointegration.* Oxford: Oxford University Press, 267-276.

Pesaran, M. H. and A. Timmerman (1995): “Predictability of stock returns: Robustness and economic significance”, *Journal of Finance,* 50, 1201-1228.

Pesaran, M. H. and A. Timmerman (2000): “A recursive modelling approach to predicting UK stock returns”, *Economic Journal*, 110, 159-191.

Plummer, T. (1989), *Forecasting financial markets: The truth behind technical analysis*. London: Kogan Page.

Pring, M. J. (1991): *Technical Analysis Explained : The Successful Investor's Guide to Spotting Investment Trends and Turning Points*. New York: Macgraw-Hill.

Ramsey, J. B. (1990): “Economic and financial data as nonlinear processes, in Gerald P. Dwyer and R. W. Hafer, eds., *The Stock Market: Bubbles, Volatility, and Chaos*. Boston, MA: Kluwer Academic, 81-134.

Satchell, S. and A. Timmermann (1995): “An assessment of the economic value of non-linear foreign exchange rate forecasts”. *Journal of Forecasting*, 14, 477-497.

Scheinkman, J. and B. LeBaron (1989): “Nonlinear dynamics and stock returns”. *Journal of Business*, 62, 311-337.

Sweeney, R. J. (1988): “Some new filter rule tests: Methods and results”. *Journal of Financial and Quantitative Analysis*, 23, 285-300.

Taylor, M. P. and H. Allen (1992): "The use of technical analysis in the foreign exchange market". *Journal of International Money and Finance*, 11, 304-314.

Table 1: Statistical properties of the NYSE Composite Index

	Level	Returns
A. Summary statistics		
Mean	190.4879	0.0454
Maximum	95.8950	0.0400
Minimum	677.5800	29.4200
Std. Deviation	32.8900	-33.1200
Skewness	185.4893	2.7446
Kurtosis	1.3399	-0.5370
Jarque-Bera	3.4901[a]	25.6498[a]
B. Autocorrelation coefficients		
1	1.000	0.050
2	0.999	-0.036
3	0.999	-0.027
4	0.999	0.000
5	0.999	-0.035
6	0.998	-0.025
7	0.998	-0.041
8	0.998	0.012
9	0.998	0.010
10	0.997	0.009
C. Summary measures of autocorrelation		
Adjusted Box Q-statistic (on 10 lags)*	92926[a]	77.289[a]
Augmented Dickey-Fuller Test**	1.0927	-45.5832[a]

Notes: [a] y [b] denote significance at the 1% and 5% levels, respectively
* This is distributed as chi-squared with l degrees of freedom, where ℓ is the number of lags. The critical value at the 0.05 significance level is 18.31
** In this one-sided t-test, the critical values at the 1% and 5% significance levels are –3.43 and –2.86 (MacKinnon, 1991)

Table 2: BDS tests for independence of the NYSE Composite Index

(A) Residuals							
	m=2	m=3	m=4	**M=5**	m=6	m=7	m=8
$\varepsilon = 0.75\sigma$	54.87305[a]	68.41499[a]	77.55682[a]	87.28173[a]	98.79450[a]	112.68815[a]	129.77594[a]
$\varepsilon = \sigma$	51.90813[a]	64.30134[a]	71.32241[a]	77.95172[a]	84.91187[a]	92.62240[a]	101.43772[a]
$\varepsilon = 1.25\sigma$	47.93073[a]	59.89205[a]	66.19658[a]	71.57854[a]	76.59581[a]	81.80406[a]	87.47073[a]
$\varepsilon = 1.5\sigma$	43.73106[a]	55.12333[a]	60.96430[a]	65.47817[a]	69.25381[a]	72.93456[a]	76.92694[a]
$\varepsilon = 1.75\sigma$	40.30905[a]	51.10630[a]	56.44100[a]	60.41696[a]	63.42536[a]	66.19201[a]	69.11901[a]

(B) Shuffled residuals							
	m=2	m=3	m=4	m=5	m=6	m=7	m=8
$\varepsilon = 0.75\sigma$	-0.29024	-0.81592	-1.41454	-1.82741	-1.63650	-1.34916	-1.15107
$\varepsilon = \sigma$	-0.24507	-0.66396	-0.69589	-0.62046	-0.36024	-0.34230	-0.19427
$\varepsilon = 1.25\sigma$	-1.12978	-1.20190	-0.88240	-0.89113	-0.75880	-0.68606	-0.77673
$\varepsilon = 1.5\sigma$	0.22256	0.45753	0.62651	0.60224	0.61697	0.45804	0.56769
$\varepsilon = 1.75\sigma$	0.52125	0.52729	0.53500	0.72718	0.83762	1.15080	1.29430

Notes: The BDS statistic is applied to the AR (6) residuals of the original series in Panel (A) and to shuffled AR(6) residual in Panel (B). We report the results for several values of the embedding dimension (*m*), and the distance (ε). The latter is related to the standard deviation of the data (σ). BDS statistics are distribute *N*(0,1) under null hypothesis of iid residual. "a", "b" and "c" denote significance at the 1,5 and 10% levels, respectively.

Table 3: GA selected values for the NN parameters

Prediction period	Embedding dimension (m)	Number of closest points (k)[a]
1997	1	1%
1998	10	0.9%
1999	5	7%
2000	3	4.5%
2001	1	7.9%
2002	3	3.6%

NOTE: [A] IN PERCENTAGE OF THE SAMPLE

Figure 1: NYSE Composite Index (1966-2002)

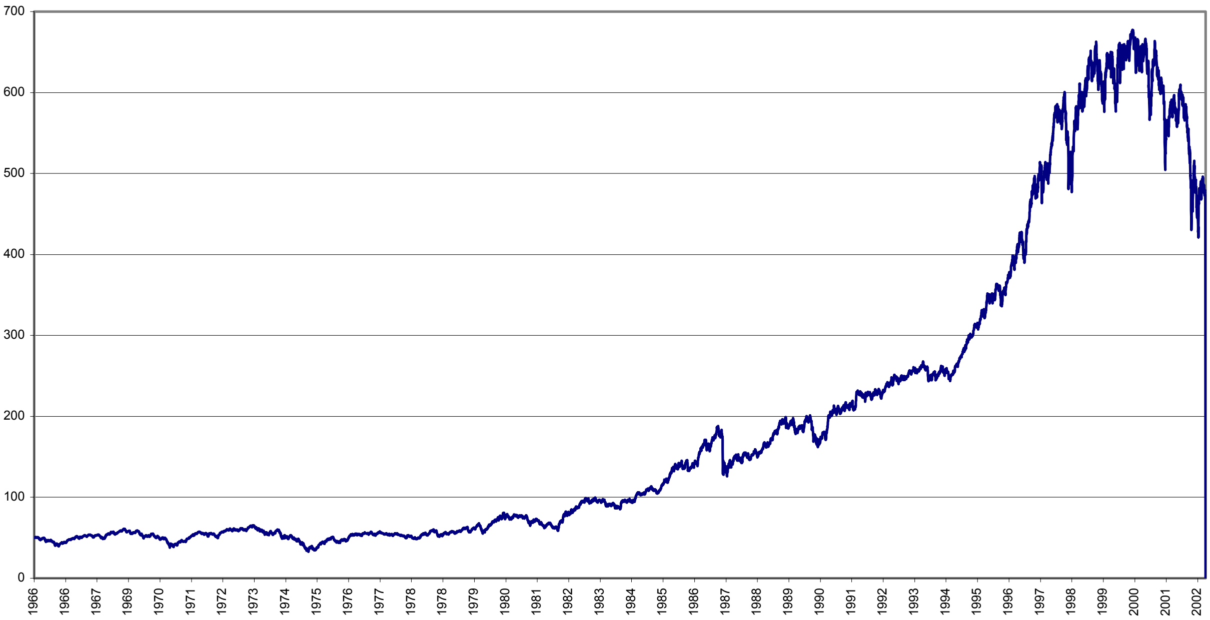

Figure 2: NYSE Composite Index (1997)

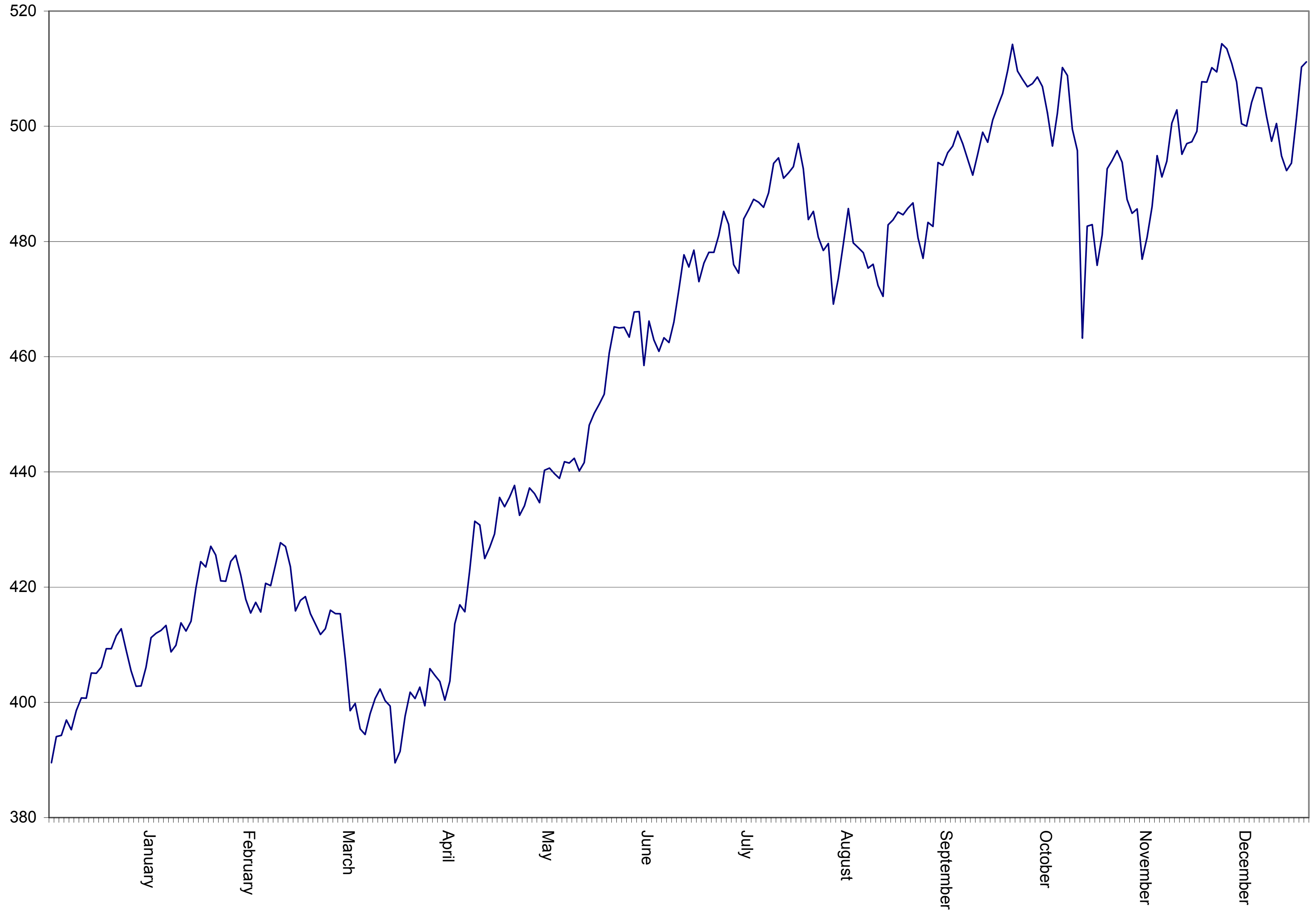

Figure 3: NYSE Composite Index (1998)

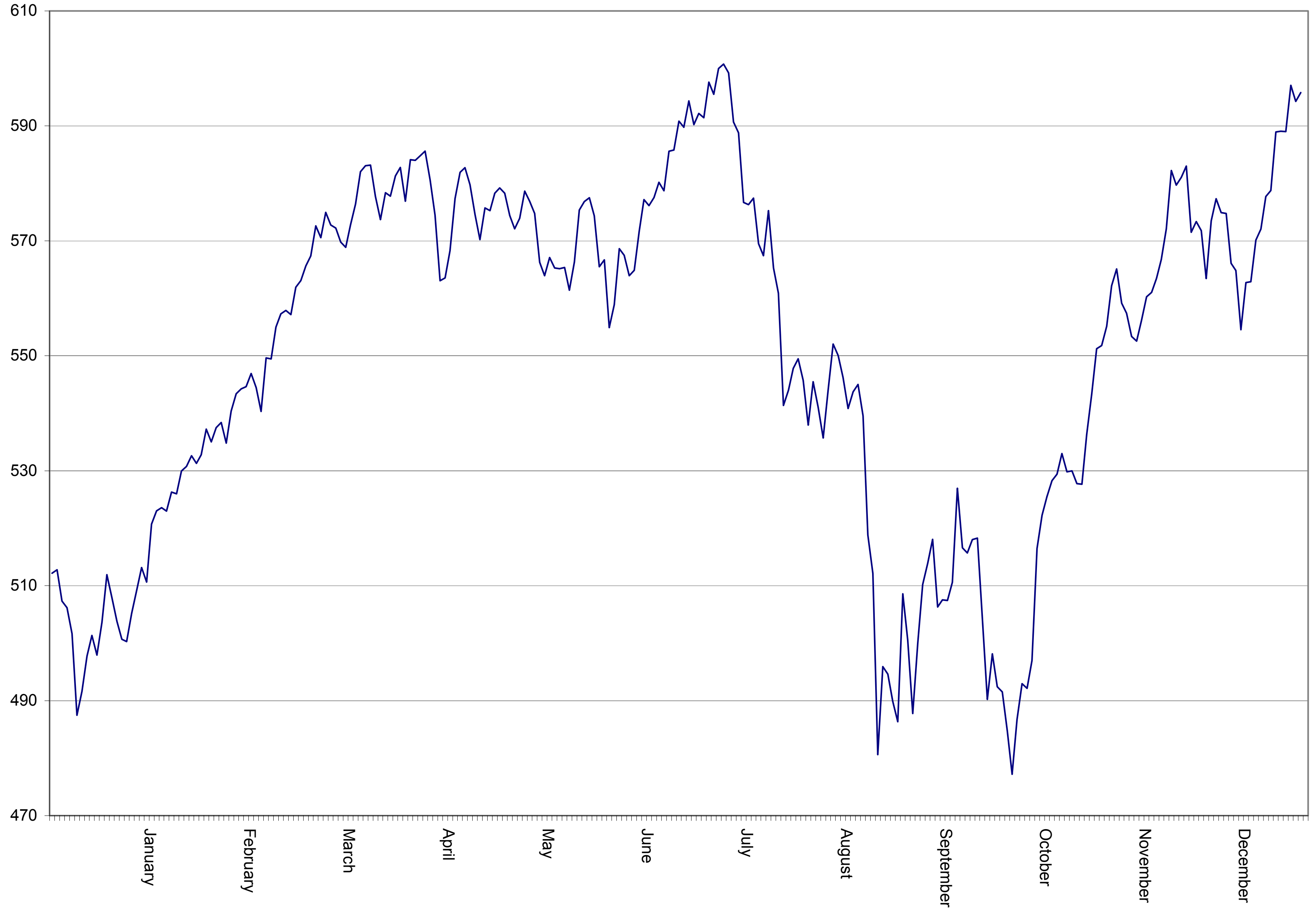

Figure 4: NYSE Composite Index (1999)

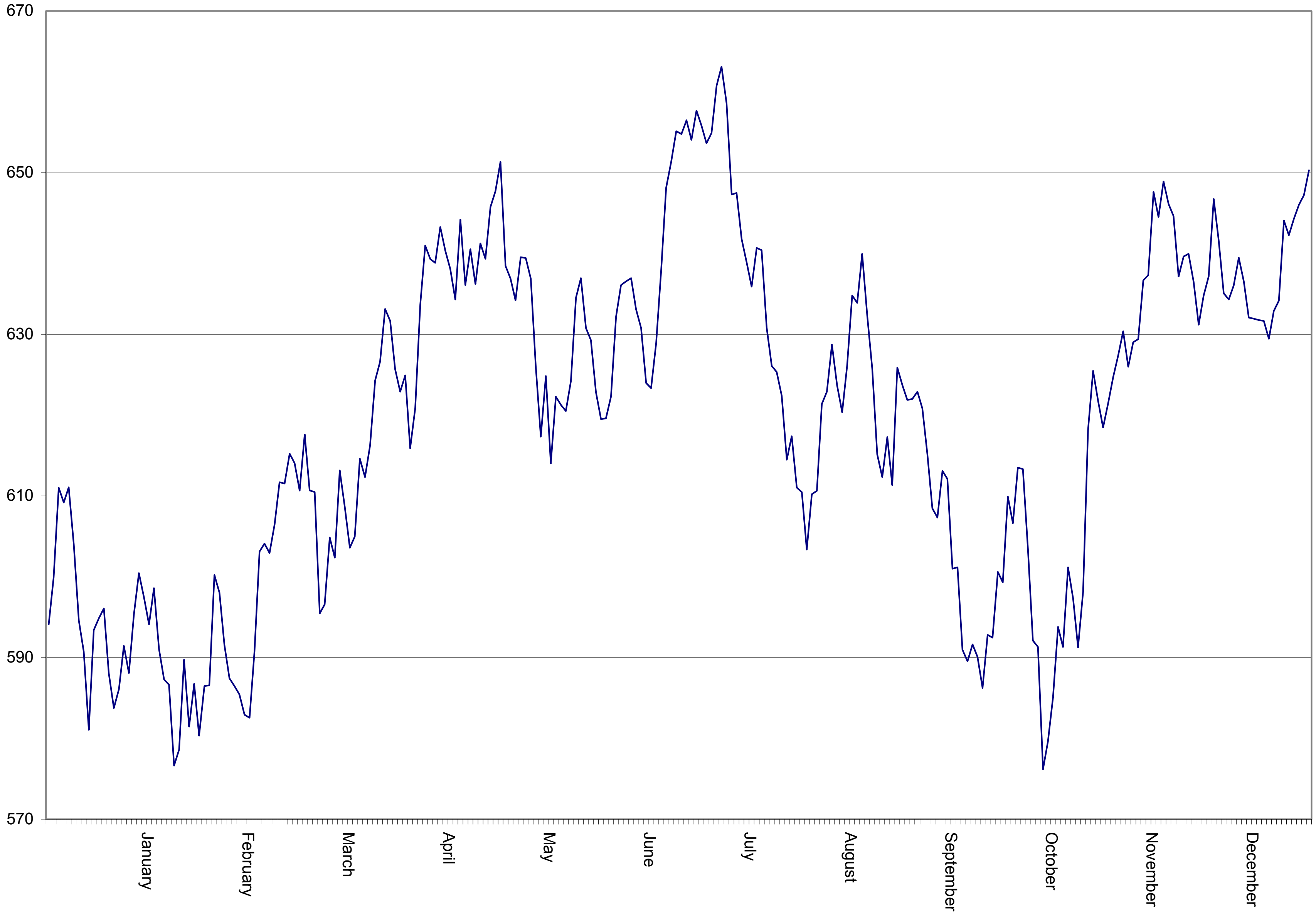

Figure 5: NYSE Composite Index (2000)

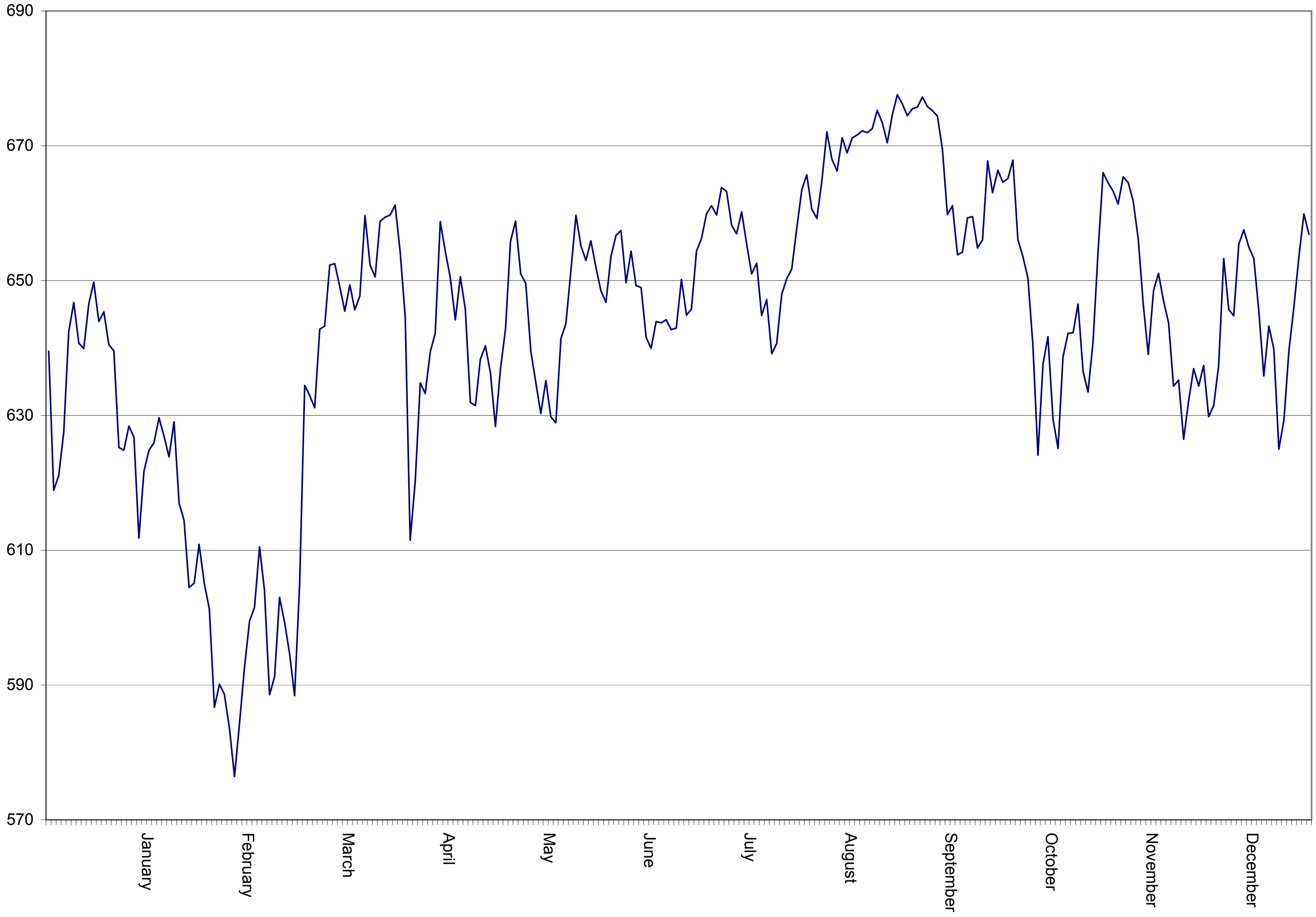

Figure 6: NYSE Composite Index (2001)

Figure 7: NYSE Composite Index (2002)

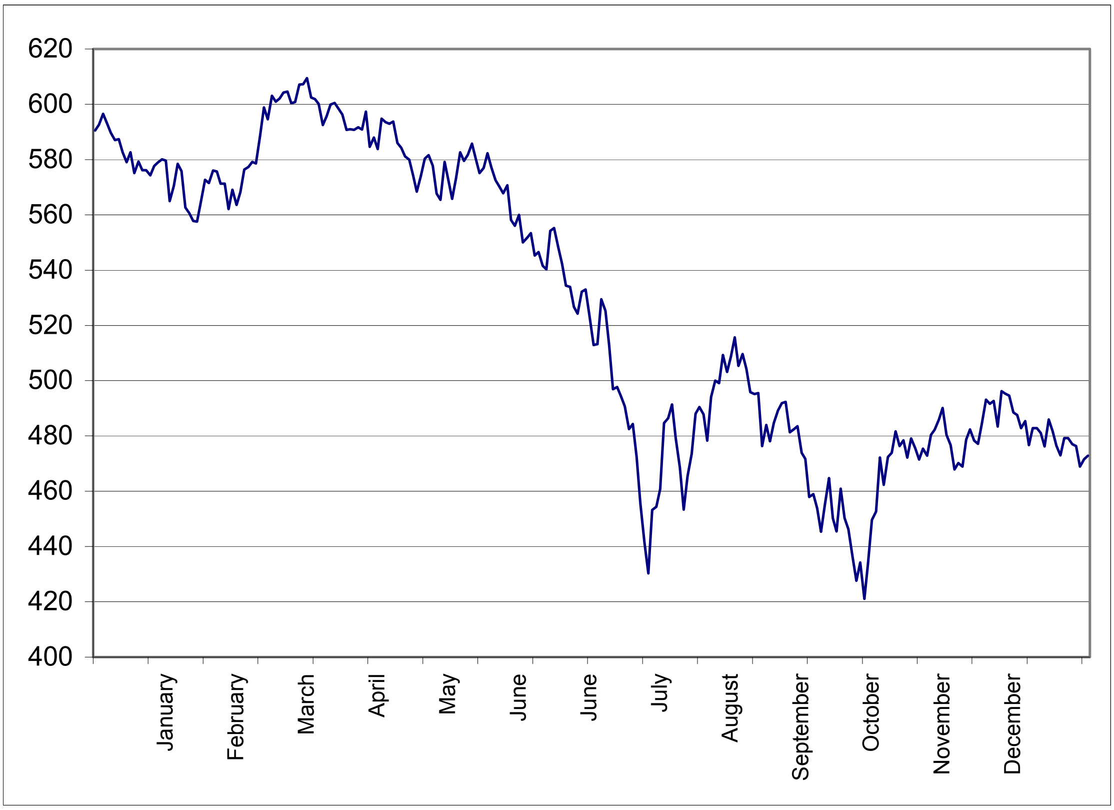

Figure 8: Out–of-sample net returns:
NN versus ARIMA, Buy-and-hold and Risk-adjusted Buy-and-hold (1997)

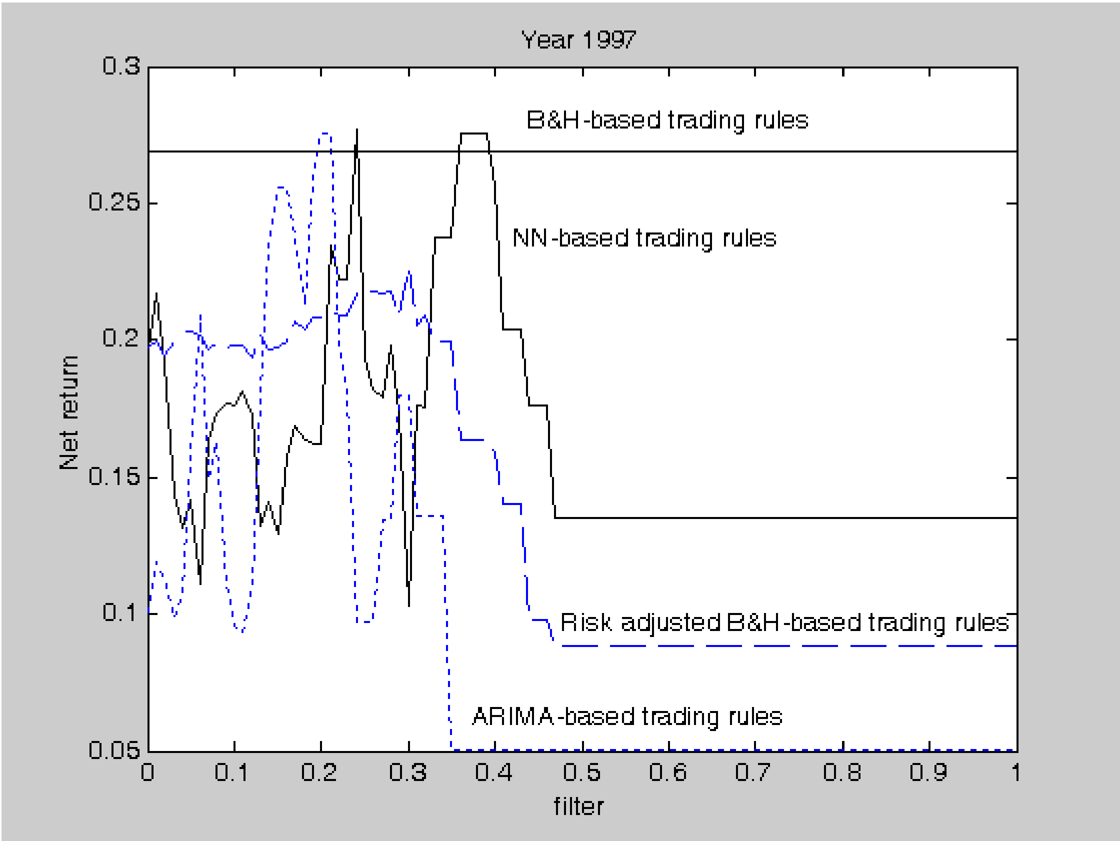

Figure 9: Out–of-sample net returns:
NN versus ARIMA, Buy-and-hold and Risk-adjusted Buy-and-hold (1998)

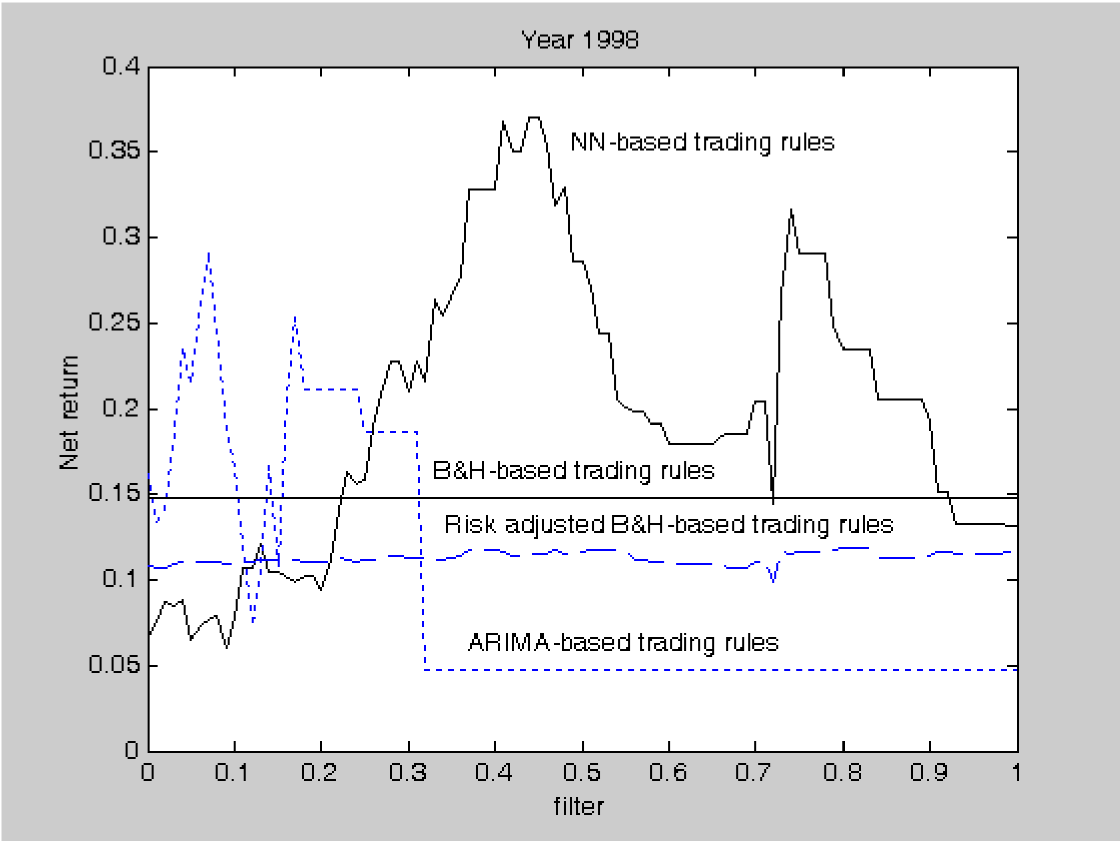

Figure 10: Out–of-sample net returns:
NN versus ARIMA, Buy-and-hold and Risk-adjusted Buy-and-hold (1999)

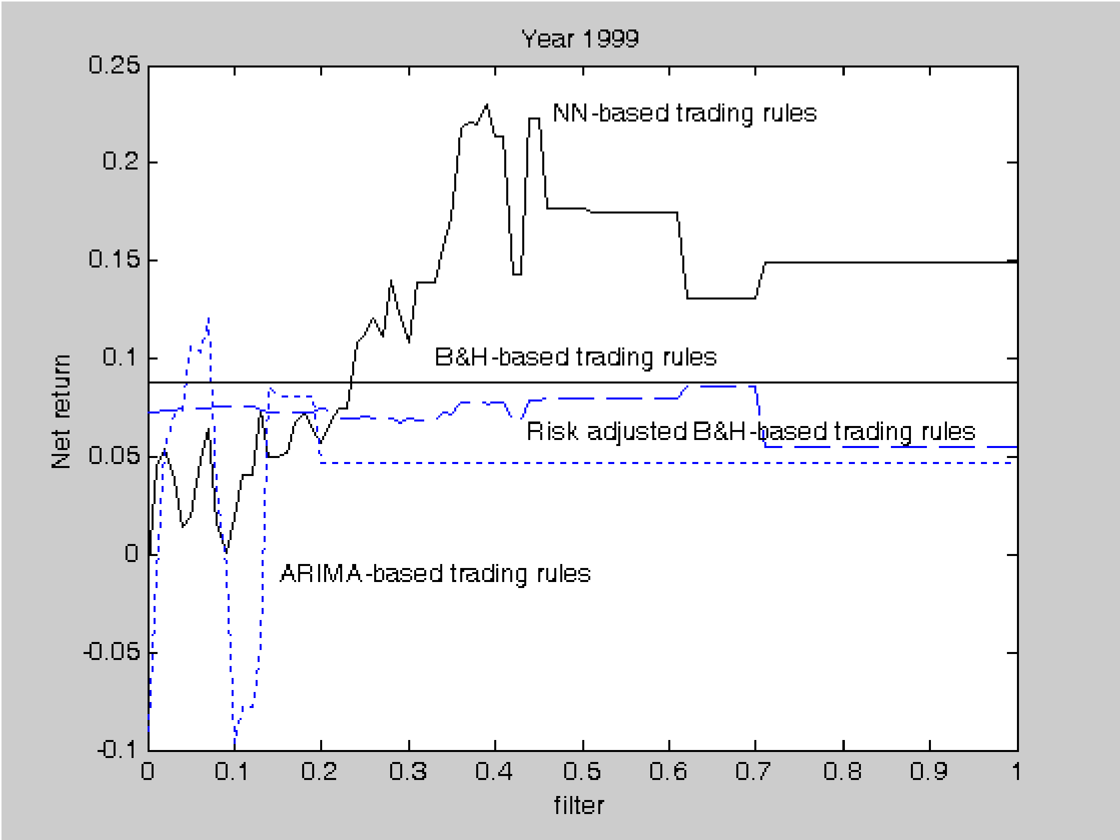

**Figure 11: Out–of-sample net returns:
NN versus ARIMA, Buy-and-hold and Risk-adjusted Buy-and-hold (2000)**

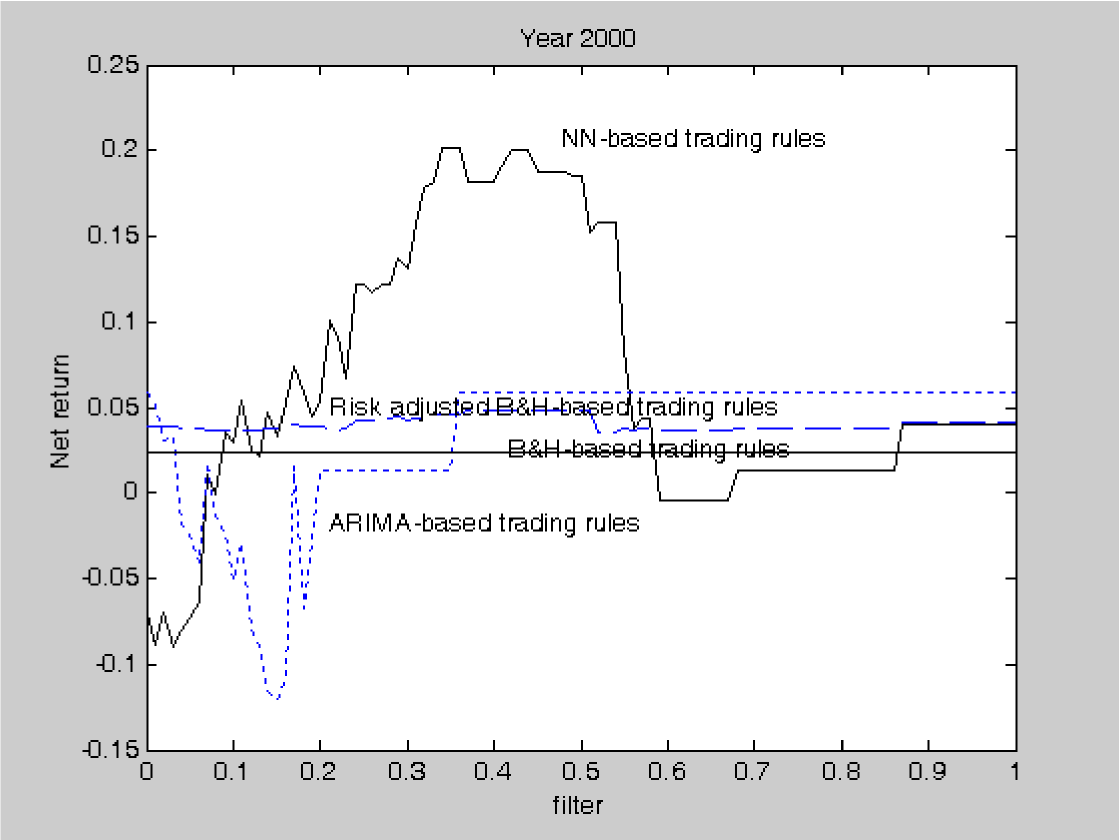

Figure 12: Out–of-sample net returns:
NN versus ARIMA, Buy-and-hold and Risk-adjusted Buy-and-hold (2001)

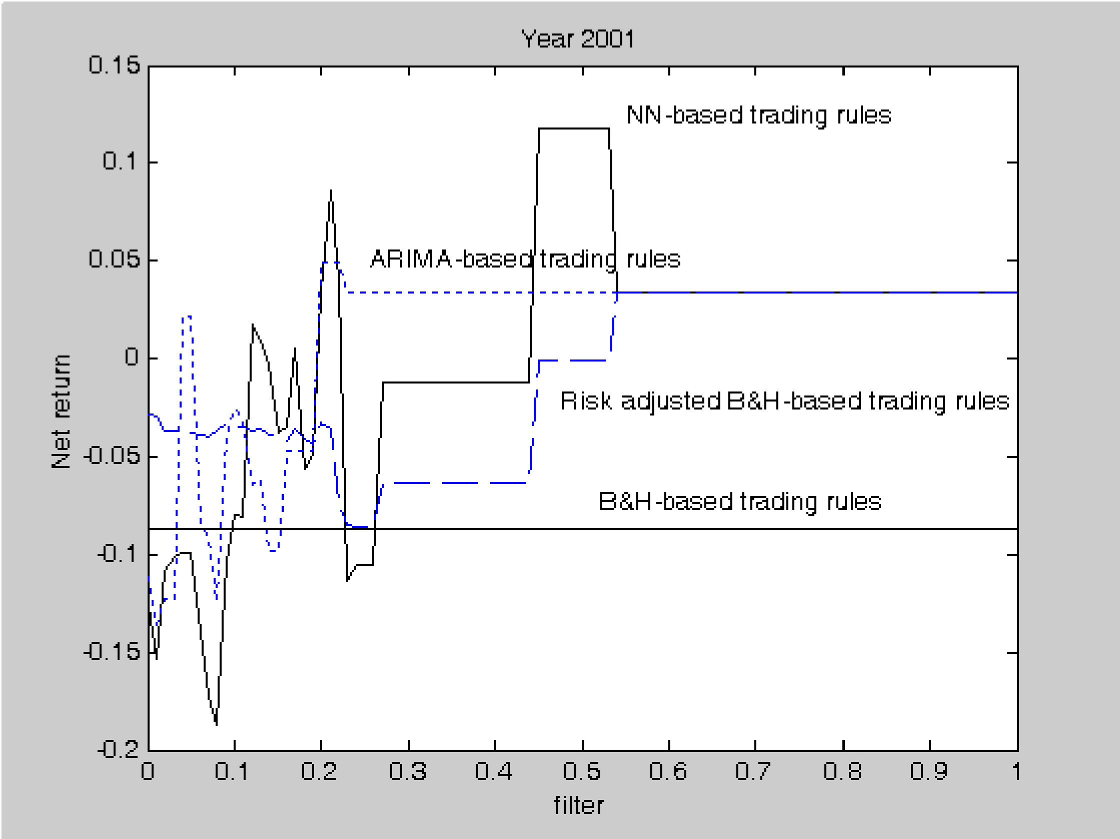

Figure 13: Out–of-sample net returns:
NN versus ARIMA, Buy-and-hold and Risk-adjusted Buy-and-hold (2002)

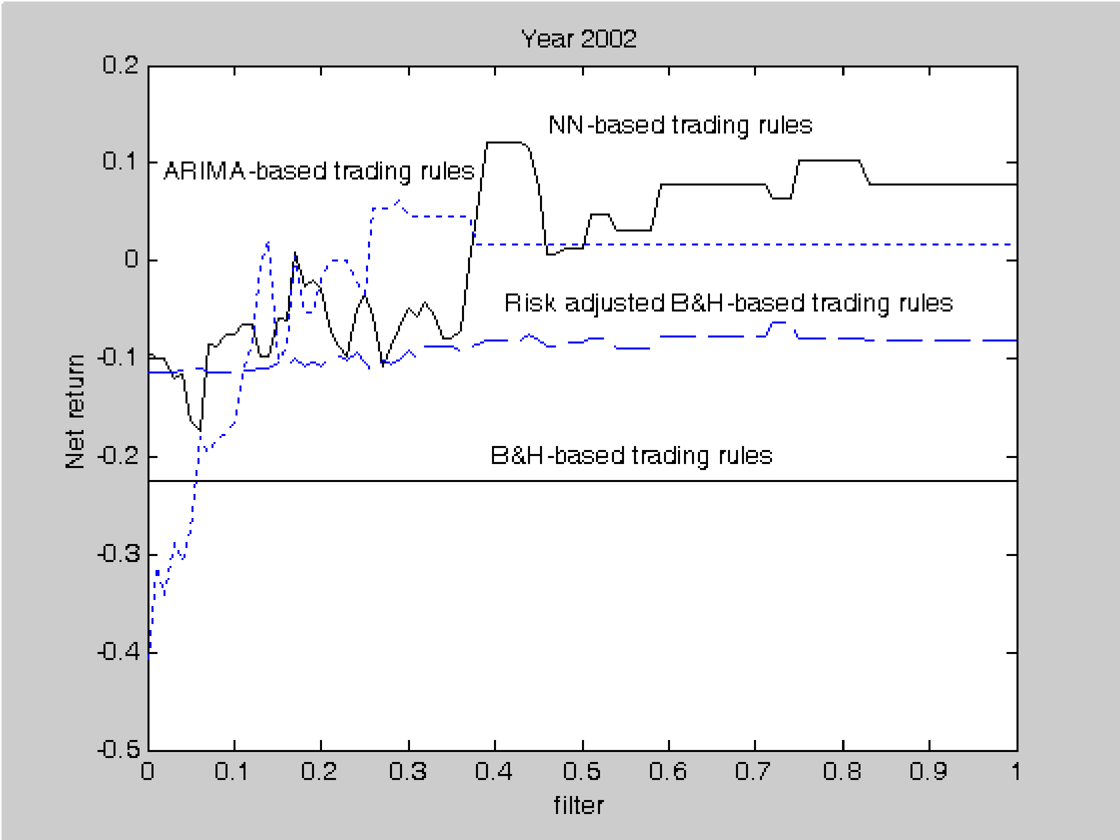

Figure 14: Sharpe ratios: NN versus ARIMA (1997)

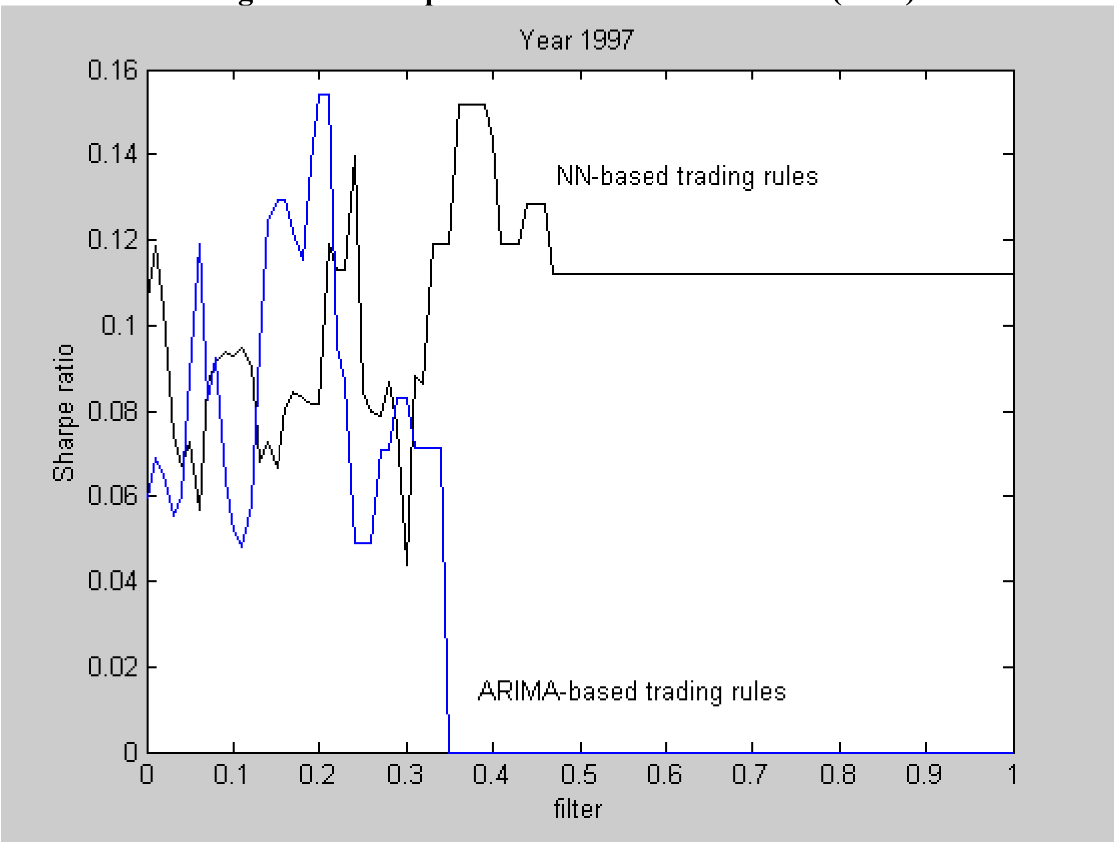

Figure 15: Sharpe ratios: NN versus ARIMA (1998)

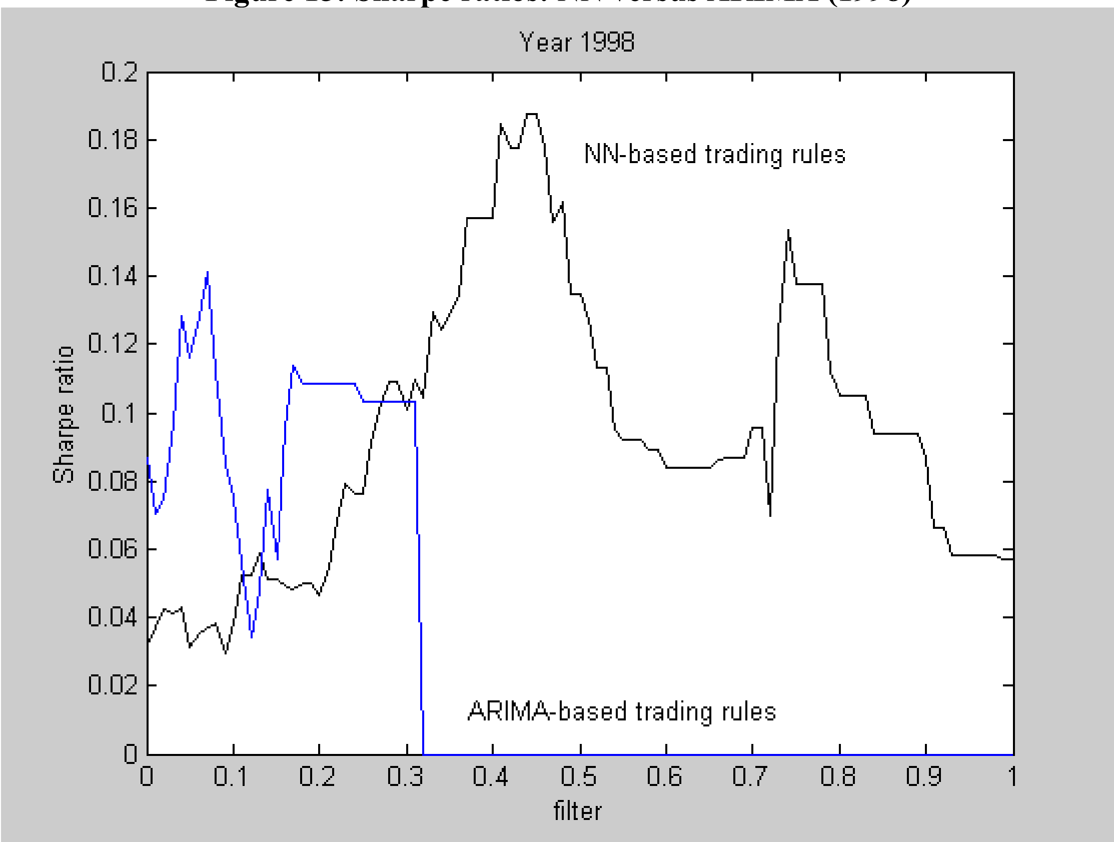

Figure 16: Sharpe ratios: NN versus ARIMA (1999)

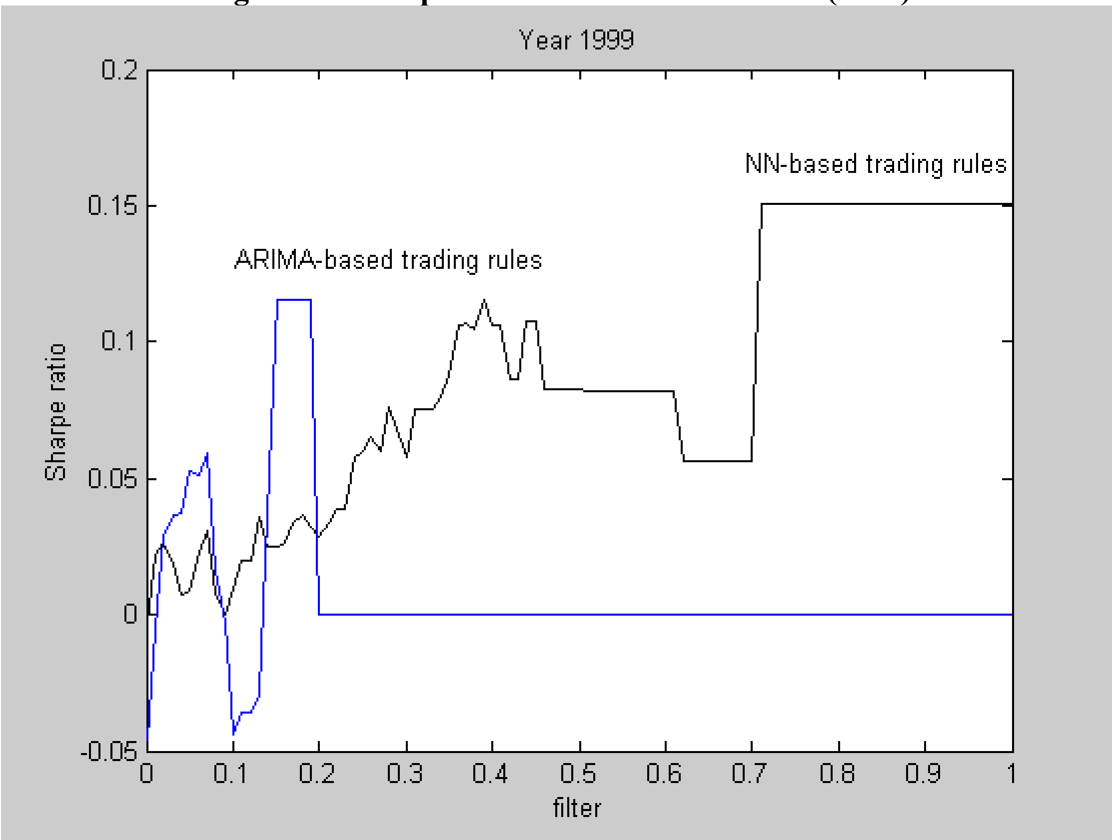

Figure 17: Sharpe ratios: NN versus ARIMA (2000)

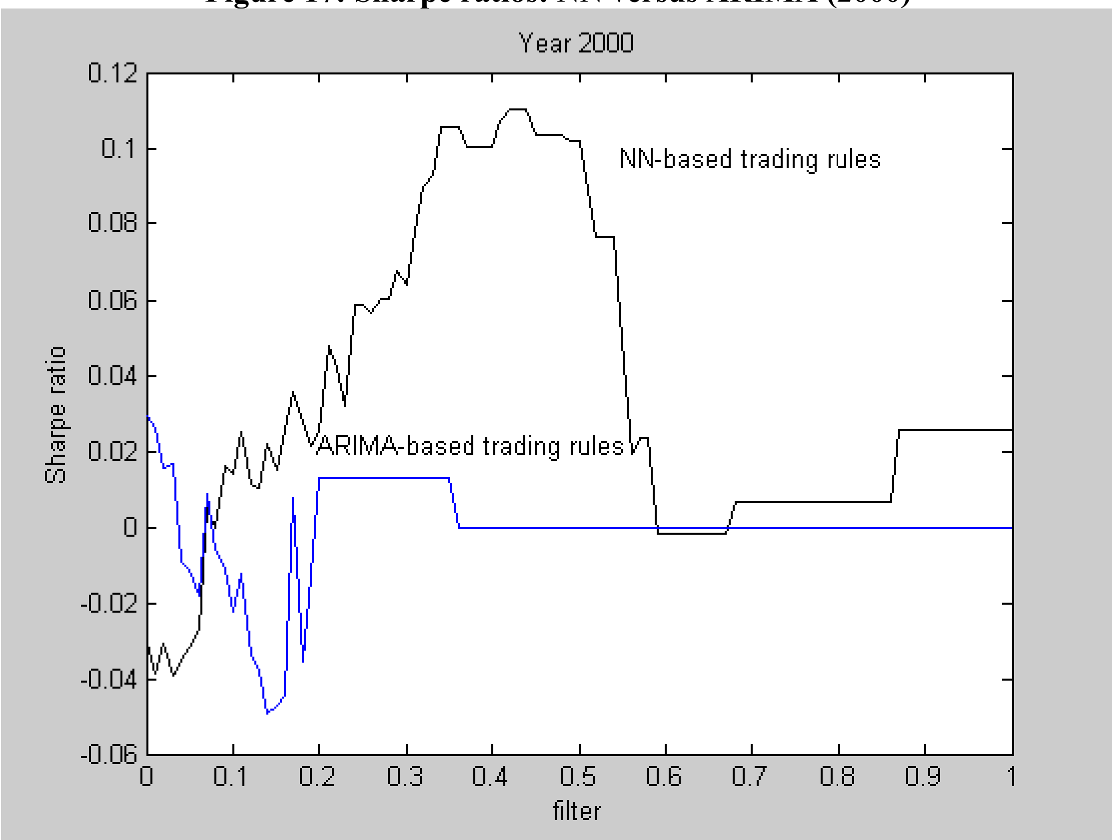

Figure 18: Sharpe ratios: NN versus ARIMA (2001)

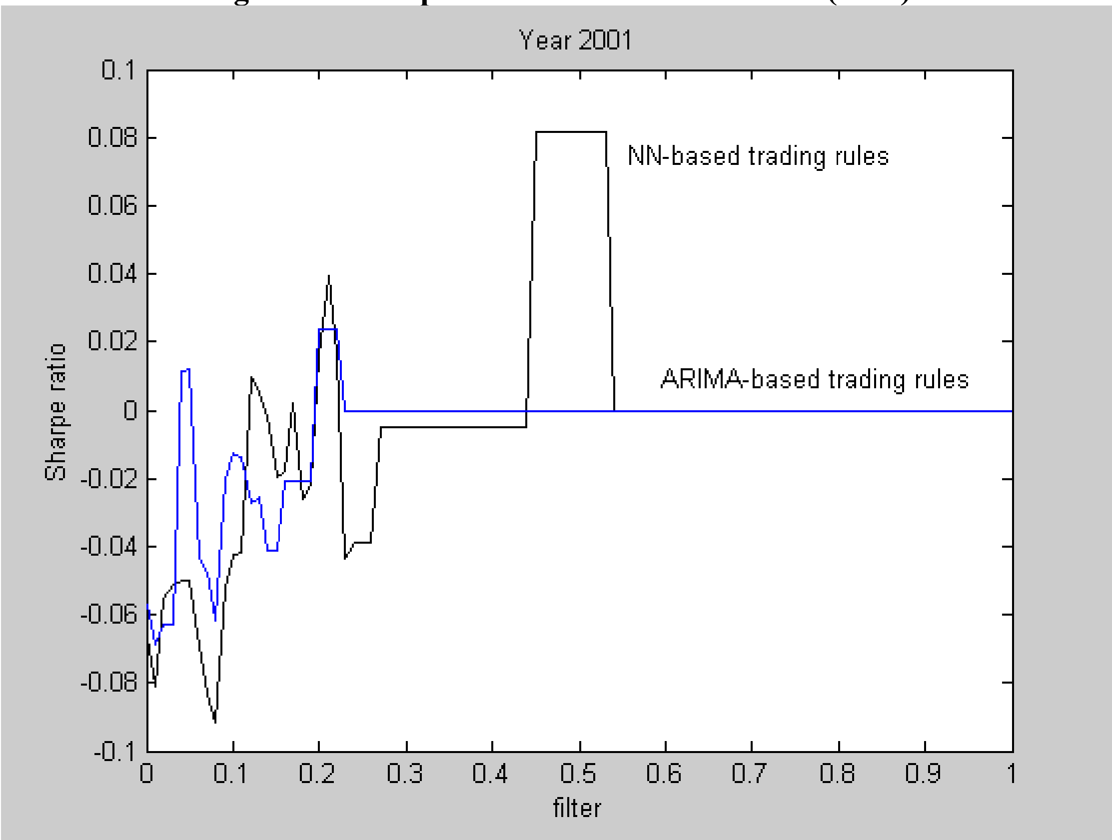

Figure 19: Sharpe ratios: NN versus ARIMA (2002)