DOES ADAPTIVE EPS FORECASTING MAKE ANALYSTS' FORECASTS REDUNDANT? *

Dmitri Kantsyrev

University of Southern California
3620 S. Vermont Ave, KAP 300
Los Angeles, CA 90089
kantsyre@usc.edu

Revised: July, 2005

*I thank Michael Magill and Fernando Zapatero for their helpful discussions and suggestions. I am grateful for the insightful comments from Jaksa Cvitanic, Christopher Jones, Cheng Hsiao, and Lloyd Levitin. The author also thanks Thomson Financial for providing the I/B/E/S database, as part of an academic program to encourage earnings expectation research.

Abstract

In this study, I examine the relative accuracy of financial analysts' and adaptive time-series earnings forecasts made at the beginning of a fiscal year. I consider IBES consensus forecasts and employ a novel forecasting approach: artificial neural networks. The central question is whether financial analysts efficiently utilize available information and produce forecasts that are more accurate than predictions of statistical models. In contrast to the existing literature, which analyzes non-adaptive forecasting techniques, I present evidence of the superiority of adaptive time-series models forecasts over financial analysts' forecasts made at the beginning of a fiscal year for a specific subset of firms. The study shows a way of differentiating companies according to statistical characteristics of their earnings, and as a result, to the relative accuracy of analysts' forecasts. I find that the relative accuracy of financial analysts' forecasts decreases with the variation of change in earnings and the forecast horizon. The evidence presented contributes to the understanding of the formation and value of analysts' predictions.

Keywords: Performance of Financial Analysts, Earnings Forecasts, Artificial Neural Networks.

I. Introduction

While earnings are the basic accounting-based measure of a firm's performance, earnings expectations are one of the strongest signals about its future prospects. Over the years, two methods of earnings predictions have been exploited: the use of financial analysts and non-adaptive time-series models. The latter are often simple statistical techniques, whereas financial analysts are viewed as a more reliable source of forecasts for all companies and at all forecast horizons. It is frequently linked to an informational advantage over time-series models that only exploit histories of earnings. On the other hand, financial analysts may not always issue objective forecasts for a number of reasons. Their forecasts may be influenced by personal career concerns or by incentive problems[1]. However, there have been few attempts to find alternative methods of forecasting in the literature since the 1980s, when the view of analysts' superiority prevailed. Therefore, it seems natural to come back to the issue of forecasting accuracy from the current perspective and to compare the relative accuracy of financial analysts' forecasts to adaptive time-series models predictions.

In this study, I consider IBES consensus earnings forecasts for the 1993-2002 period and employ adaptive forecasting techniques, in particular, a novel approach: artificial neural networks. Neural networks can detect systematic patterns, learn and adapt to underlying relationships. They are data driven and therefore useful where one does not have particular beliefs about functional forms. I provide tests of the relative accuracy of financial analysts' earnings forecasts by considering rank orders and the direction of change measure. It allows me to identify models that obtain a better forecast accuracy for the greater number of companies and to answer the question of which models have a better ability to recognize the sign of future changes in earnings.

I demonstrate the importance of statistical characteristics of a firm's earnings for assessing the relative accuracy of alternative forecasting methods. I find evidence of the superiority of adaptive time-series models forecasts over financial analysts' forecasts made at the beginning of a fiscal year for companies with highly volatile earnings. This suggests that financial analysts mainly predict the overall market behavior and have a lack of ability to predict firm specific fluctuations. On the other hand, the relative accuracy of adaptive time-series models forecasts increases with the variation of change in earnings and the forecast horizon. It is apparently caused by superior abilities of artificial neural networks to determine nonlinear systematic patterns in volatile earnings

---

[1] Research on systematic errors in analysts' earnings forecasts has produced a diverse set of incentive-based explanations intended to account for them. Francis and Philbrick (1993) find that analysts incorporate optimism into their forecasts to repair management relationships, following sell recommendations. Lin and McNichols (1998) find that co-underwriter analysts' earnings forecasts are more favorable than those made by unaffiliated analysts. Dugar and Nathan (1995) find that analysts exhibit greater optimism for firms that are investment-banking clients. McNichols and O'Brien (1997) suggest that the observed bias is a result of the selection process when analysts with relatively unfavorable information decide to exit the pool of forecasters. H. Hong *et al*. (2000) provide evidence that the analysts' behavior is consistent with career-concern-motivated herding theories.

and to recognize downward moves in the environment of generally rising earnings. In addition, I find that financial analysts produce less accurate two-year-ahead forecasts made at the beginning of a fiscal year than any other adaptive time-series model. It leaves in question the existence of analysts' forecasts as a reliable measure of a firm's expected performance at the beginning of a fiscal year.

Next, I examine whether additional information processed by analysts is constructive or if it is a noise, which hinders the discovery of systematic patterns. I consider the relative informational content of forecasts and show that time-series models have strength of their own. They often contain information missing in analysts' predictions. Furthermore, for a specific subset of firms, a forecast horizon and a time-series model, I demonstrate that all information contained in analysts' predictions is already included in a time-series model forecasts indicating that having more information is not necessarily useful. This result suggests that financial analysts either underestimate the importance of information contained in histories of earnings or cannot properly filter the extensive set of all available information. Finally, I study the relationship between the number of analysts issuing forecasts for a specific company, the standard deviation of individual forecasts and the relative accuracy of financial analysts' consensus forecasts. I show that not the size, but the type of the company is a main determinant of the financial analysts' relative forecast accuracy. It supports the finding that more information is not necessarily beneficial for the accuracy of predictions.

Underlying this research is abundant theoretical and empirical literature on the subject of time-series of earnings that has received wide attention since 1973, when the Security and Exchange Commission announced its intention to require management forecasts to be made public. Earlier studies, such as Ball and Watts (1972), Albrecht *et al.* (1977), Watts and Leftwich (1977), argue that histories of past annual earnings per share contain almost no information about future earnings, and conclude that earnings are best described as random processes. On the other hand, works by Brown and Roseff (1979), Collins and Hopwood (1980), Hopwood *et al*. (1982) consider quarterly earnings as inputs to forecasting models and state that quarterly EPS have appeared to yield the predictions of future annual earnings that often compete in accuracy with the random walk model.

Empirical tests comparing the accuracy of financial analysts' earnings forecasts to the accuracy of non-adaptive time-series models predictions claim analysts' superiority. Brown and Roseff (1978) analyze fifty firms followed by a single analyst, Value Line Investment Survey, and provide evidence of Value Line's superiority over the Box and Jenkins and naive models. Fried and Givoly (1982) note that the broadness of the information set employed by analysts and their reliance on information released after the end of a fiscal year appear to be important contributing factors to the analysts' superior performance. Similarly, Brown *et al.* (1987) attribute the analysts' superiority to a timing advantage and an information advantage[2]. They also show a positive association between

[2] Timing advantage - more information is available after the earnings announcement; information advantage - more information is used by analysts than historical earnings.

the firm size and the advantage of financial analysts' forecasts over time-series based forecasts.

However, note that most empirical works have studied the accuracy of non-adaptive statistical models, specifications of which are fixed through time. This approach neglects the changing nature of data generating processes and may not provide the most accurate forecasts. My approach is to consider the use of *adaptive* time-series models as a tool for forecasting earnings. In contrast to non-adaptive statistical models, the main feature of this method is the assumption that the underlying relationship between past and future earnings may be evolving over time. As approximate means by which I hope to capture this phenomenon, I not only re-estimate parameters of statistical models, but also choose a new specification each time new data become available.

The organization of this paper is as follows. Section II describes data used and a sample selection process. It also illustrates differences between the suggested adaptive statistical approach and the non-adaptive one that has been continuously exploited in the past. Section III discusses estimation methodology. The empirical findings are presented in Section IV, while the final Section V states conclusions drawn from this work.

II. Experimental Design

A. *Data and Sample Selection*

I use consensus forecast data from the Institutional Brokers Estimate System Summary file. In the IBES database, consensus forecasts of firms' earnings are the means of all analysts' estimates outstanding as of the Thursday before the third Friday of each month. The choice of consensus forecasts in favor of individual analysts' forecasts is not arbitrary. Investors often rely on consensus forecasts of earnings as measures of a firm's future performance. In firm valuation models, the intrinsic value of a company also depends on consensus expectations of future earnings. In testing such models, consensus forecasts are the appropriate proxies to be used, and an ex-post accuracy is not a key motive for using consensus measures. In contrast to my approach, most of the existing studies that compare the accuracy of time-series models and analysts' earnings predictions consider single analyst's forecasts. Therefore, the analysis of consensus forecasts rather than that of individual analysts' forecasts constitutes the first distinction of the current research.

The IBES earnings forecasts database covers approximately 14,500 companies for different periods starting in 1977. I restrict attention to the December year-end firms. In addition, for a firm to be included in the sample, there should exist one-year-ahead, two-year-ahead and one-, two-, … , seven-, eight-quarters-ahead forecasts by at least one analyst in March of each year starting in 1993. Because most calendar-year-end firms announce their annual earnings between January and March, I use analysts' forecasts from the month of March. It ensures that analysts' as well as statistical models forecasts are conditioned on the same knowledge of previous years earnings. Thus, I control for the timing advantage, while the information advantage becomes the subject for the test.

The selection procedure results in a sample of firms with ten-year histories of annual and quarterly forecasts. I choose the ten-year forecast comparison window because of the following consideration: time-series models in this study require histories of actual earnings to estimate parameters of the models and to generate forecasts. Therefore, there is a tradeoff between the number of years used to compare forecast errors and a selection bias, which increases with extending histories of actual earnings farther in the past. Note that the common criticism of most published studies is that, by dealing with only a few forecast comparison dates, they report results that may be specific to relatively short time intervals. For example, Brown and Roseff (1978) compare forecasts for only four years. The relatively longer forecast comparison period is the second distinction of this paper.

To finalize the selection procedure, I randomly choose forty-eight firms from the set of companies that satisfy all requirements discussed above. I admit that it is possible that there exist some sample bias due to the limited coverage of firms by financial analysts. This bias is towards a greater coverage of large and somewhat older firms that have forecast data reported by the IBES in March for ten consecutive years. For this reason, extrapolations to larger populations should be made with care.

I take actual earnings (1973-2002) from the Compustat database. Note that both analysts' forecasts and actual earnings series are Earnings per Share before Extraordinary Items and Earnings from Operations as soon as the latter became available around 1986. I adjust all data for stock splits and stock dividends.

B. *Time-Series Models.*

In this section, I discuss the adaptive time-series models that are used to forecast future earnings. By an adaptive model I mean that a new specification is chosen before each new rolling forecast is constructed. The notion of *adaptability* or *real-time* forecasting constitutes a key difference between the current research and the existing literature, which conditions the analysis on a fixed, non-adaptive forecasting model assumed to be in effect throughout an entire sample period. The latter approach misses the important detail that a time-series model, which reasonably describes an earnings-generating process in one period, may be inappropriate in another. It can be caused by changes in the macroeconomic situation, changes in factors affecting the industry or the firm. For example, changes in a firm's earnings may be attributed to general economic factors such as cyclical changes, which are represented by economy-wide ups and downs caused by a business cycle, or to structural changes like changes in the demographic structure or technologies. Shifts in the industry's position as well as transformations in the political situation may change a company's fortune over time. To illustrate the idea, I present earnings per share of two companies in Figure 1. In the first graph, earnings are highly cyclical, indicating that the company's performance is very sensitive to the overall market condition. Whereas in the second graph, earnings follow the market quite closely, suggesting that we are dealing with a *typical market* firm. In addition, note that there is a drastic change in the market earnings pattern around 1992.

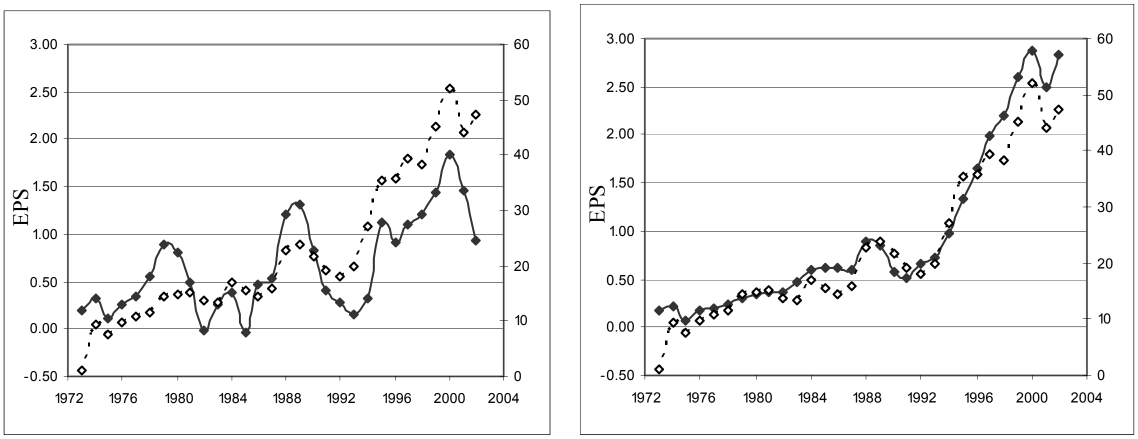

Figure 1. EPS before Extraordinary Items (1973-1985) / EPS from Operations (1986-2002) of *Alcoa Inc.* and *Avery Dennison* (presented by the solid line, left scale), and *S&P500* (presented by the dotted line, right scale).

By estimating adaptive linear and non-linear models, I address the following question: "Is there evidence that adaptive models are valuable in forecasting earnings?" If so, we have a clear alternative to often expensive financial analysts' earnings forecasts. Swanson and White (1995, 1997) find that such models are useful when the variable of interest is the spot-forward rate differential: they show that adaptive linear vector autoregression models often outperform professionally available survey predictions, as well as no-change and non-adaptive linear models of key macroeconomic variables.

The class of *non-linear* time-series models is presented by artificial neural networks that are known to be universal function approximators and are capable of exploiting non-linear relationships between variables[3]. Neural Networks are applied across a wide range of disciplines: medicine, engineering, geology, and physics. In contrast, for many years, linear modeling has been a commonly used technique in economics and finance since linear models have well-known optimization strategies. Where the linear approximation was not valid, the models suffered accordingly. Only recently, artificial neural networks became the focus of attention as a possible vehicle for forecasting economic and financial variables. Kuan *et al.* (1995) consider exchange rate forecasting and conclude that neural networks have significant market timing ability and significantly lower out-of-sample mean square prediction error relative to the random walk model. Tkacz (2001) finds that neural networks yield statistically lower forecasts errors for the growth rate of real Canadian GDP relative to linear models.

An artificial neural network is a sophisticated information processing technique that is inspired by the way the human brain processes information. The major element of this mechanism is a novel structure of the information processing system that is composed of highly interconnected processing elements. These elements, or units, are organized in layers. It is customary to distinguish the input layer, which supplies input data, hidden layers, and the output layer. The greater the number of hidden layers, the

[3] For further discussions see, for example, Bishop (1995), Fausett (1994), Hornik *et al.* (1989).

greater the complexity of the system, and as a result, more cases are required to estimate the model. Due to the small number of cases available in this study, the networks I consider contain only one hidden layer and, therefore, can be represented by a simple functional form:

$$\varphi_h(\omega_j, \omega_{ji}, \xi_j, \xi_{out}) = \psi\left(\sum_{j=1}^{h} \omega_j \left[\psi\left(\sum_{i}^{n} \omega_{ji} a_i + \xi_j\right)\right] + \xi_{out}\right), \tag{1}$$

where ω_{ji} denotes the weight for the connection between input i (total n inputs) and the processing unit j in the hidden layer (total h units in the hidden layer), ω_j denotes the weight between unit j in the hidden layer and the output unit, ξ_j and ξ_{out} are the threshold values and ψ is a given non-linear activation function; in this case, it is the logistic cumulative distribution function $\psi(z) = 1/(1+\exp(-z))$.

The network interpretation of Equation (1) is as follows. The input units send signals $(a_1, \ldots, a_n)$, which represent historical earnings in this study, over the connections to the units in the hidden layer. Each connection can amplify or reduce the signal by weight, ω_{ji} , which controls the strength and the polarity of the relationship. The modified signals that arrive at the intermediate hidden units are first summed and after the addition of a threshold, ξ_j, converted to a hidden unit activation, $\psi(\,)$. The operation of the next level is similar when hidden unit activations are sent through the connections to the output unit. The output unit performs a biased weighted sum of its inputs and passes the activation level through the transfer function to produce the output. Thus, the network has a simple interpretation as a form of input-output model with weights and thresholds as free parameters of this model. Barron (1991) demonstrates that a feedforward neural network can achieve an approximation rate $O(1/h)$ by using a number of parameters $O(hn)$ that grows linearly in h, whereas traditional polynomial and trigonometric expansions require exponentially $O(h^n)$ terms to achieve the same approximation rate. Consequently, neural networks are relatively more parsimonious than the series expansions in approximating unknown functions. This property makes neural networks an attractive econometric tool in nonparametric applications.

The neural network is data driven in that it learns only from the data presented to it and has no underlying parametric model. The greater the number of units in the hidden layers, the more the network is able to cope with non-linear relationships, but the danger of overfitting increases. The network is only trained on a training set and it is not the same as minimizing error on the error surface of the underlying and unknown model. Thus, a major flaw in the approach outlined above is that it does not minimize the error we are interested in: the error that the network will make when it encounters new and unseen cases. For this reason, some fraction of the data set must be reserved for cross-verification. The verification data are taken out from the training data and not, in fact, used for training in the back propagation. Instead, they are kept for use in an independent check on the progress of the algorithm. As the training progresses, the training error essentially drops and verification error drops as well. However, if the verification error starts to rise, it indicates that the network starts to overfit and training should cease. A larger verification set is likely to be more representative. However, it does take the data away from the training set. It is, therefore, necessary to strike a balance between the training and verification data sets. Experimentally, I find that the optimal size of the verification set is about 25-30% of the total number of available data points and that verification cases should be shifted more to the end of the data sample.

The *linear models* considered in this paper are represented by the *ARIMA(p,d,q)* specification:

$$\left(1-\phi_1 L-\phi_2 L^2-\ldots-\phi_p L^p\right)\left(1-L\right)^d a_t = c+(1+\theta_1 L+\theta_2 L^2+\ldots+\theta_q L^q)\varepsilon_t, \qquad (2)$$

where a_t denotes annual earnings per share and d is equal to zero if the earnings generating process is stationary and equal to one or two if there is evidence of nonstationarity. To find a suitable model specification and to estimate the parameters I adopt the Box and Jenkins (1970) modeling technique. First, I consider patterns of autocorrelation and partial autocorrelation functions to identify the specification of the model. Then, I estimate parameters by OLS with the Schwarz information criterion as a guide to model selection and perform diagnostic checks on residuals. I repeat the procedure every time a new data point becomes available. This technique certainly captures the spirit of real-time forecasting and, therefore, enables me to select the most appropriate linear time-series specification that is consistent with each firm's earnings generating process at a specific point in time. Consequently, forecasts obtained by this method should be superior to forecasts of ad hoc time-series models applied to all firms' time-series data.

Another linear time-series model, which plays the role of a benchmark in this study, is a random walk with drift:

$$f_t = a_{t-1}+\delta+\varepsilon_t, \qquad (3)$$

where the drift parameter, δ, is specific for each firm and period of time. I estimated it as the average earnings change from one year to the next using two to fifteen years of annual earnings data preceding the year for which a forecast is desired:

$$\delta = \frac{a_{t-1}-a_{t-n}}{n-1}, \qquad (4)$$

where n is a firm and time specific lag parameter. This is a main attribute of the proposed random walk process. It transforms the model into the framework of adaptability. To evaluate lags, I implement the following procedure. First, starting in 1987, for each company using (3) and (4) with n ranging from two to fifteen, I generate a sequence of one-year-ahead *ex post* forecasts for 1988-1992. Next, I choose the value of n that results in the smallest MSE over this period. Finally, I use (3) and (4) with the found value of n to predict earnings for 1993-1997. I repeat the procedure in 1997 to find new firm specific values of n that are used to predict 1997-2002 earnings. I find that values of n decrease with the decline in earnings volatility. If a company's earnings are steady, only recent earnings are important for forecasting, and n is small. On the contrary, if earnings are volatile, the value of n is relatively large. It tends to incorporate a relatively long-term trend in forecasting.

III. Estimation and Model Selection

A. *Estimation*

In this section, I discuss the estimation of two classes of models described above. I estimate the parameters of all non-linear and linear models using only a finite window of past data rather than all of the previously available data. By pursuing this strategy, I assume that the underlying earnings generating process may be evolving through time. I use annual and quarterly earnings data as inputs to neural networks and annual data as inputs to linear models. Accordingly, throughout the paper, by an annual/quarterly neural networks model, I mean that the model exploits and predicts annual/quarterly data. I estimate annual models using twenty years and quarterly models using eighty quarters of earnings data immediately preceding the year for which a forecast is desired, and obtain one-, two-year-ahead and one-, two-, … , seven-, eight-quarter-ahead forecasts for each firm, year and model. Then, I add quarterly forecasts and obtain a forecast of annual earnings for a given firm by a given model. I re-estimate the configuration of neural networks, the specification of the ARIMA model, and the parameters of these models each year during the period of 1993-2002.

The type of linear econometric models used and their underlying assumptions are standard. Therefore, I now turn to the discussion of non-linear neural networks estimation. In practice, there are mainly two tasks in building neural networks: a suitable network structure (the number of hidden units) must be determined, and unknown network parameters must be estimated. The main feature of neural networks is that they learn the input/output relationship through training. The training data contain examples of inputs together with the corresponding outputs, and the network learns to infer the relationship between the two. The training proceeds by back propagation developed by Rumelhert *et al.* (1986), which uses data to adjust the network weights, ω, and thresholds, ξ, so as to minimize the error in its predictions:

$$\theta_h^* = \arg\min E\left|y - \varphi_h(a, \omega, \xi)\right|^2 . \tag{5}$$

The estimation is performed through iterations. Each iteration of the training process proceeds as follows: first, the network is presented with a set of training examples from which weight and threshold adjustments are made. As a result, the training algorithm incrementally seeks for the global minimum by calculating a gradient vector of the multidimensional error surface and making a downhill move. Then, the network is tested using independent verification data to find the ability of the network to generalize on the unseen data. Training stops at the iteration where the MSE for the verification set starts to rise indicating overfitting.

The second task in practice is to establish a suitable network structure. As the activation function, ψ, can be chosen quite arbitrarily, this task reduces to determining the network complexity, i.e. the number of lagged variables, the number of hidden layers and the number of units in these layers. Although back propagation can be applied to networks with any number of layers, Cybenko (1989) shows that only one layer of hidden units suffices to approximate a large class of functions to arbitrary precision, provided that the number of hidden units, h, is adequately large and the activation functions, ψ, are non-linear. On the other hand, while a simple network (few hidden units) may not be able to approximate well, an excessively complex network (many hidden units) may overfit the data. For this reason, one should find a balance between the network complexity and the ability to predict unseen data.

For each firm and year, I estimate networks with different number of lags: three and five for the annual data and four, eight, and twelve for the quarterly data. It means that if, for example, the lag is equal to four in the case of quarterly forecasts, then the inputs to neural networks consist of quartets $(x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4})$ and the output is a single earnings number x_t. Accordingly, the first annual rolling sample consists of seventy-six inputs $(x_1, x_2, x_3, x_4), (x_2, x_3, x_4, x_5), \ldots, (x_{76}, x_{77}, x_{78}, x_{79})$ and their corresponding outputs $x_5, x_6, \ldots, x_{80}$. Thus, I produce forecasts that one could make with the model as time progresses. I also test different numbers of units in the hidden layer for each lag value[4]. I find that the neural networks perform the best with the next number of units: two units $(h=2)$ for lags equal to three and four, three units $(h=3)$ for lags equal to five, and four units $(h=4)$ for eight and twelve.

B. *Measurements of forecast accuracy.*

To assess the out-of-sample predictive abilities of alternative forecasting models, I compute the following statistics for each company and forecast horizon. The first is the mean-squared error, since it is the most frequently quoted measure in the forecasting literature:

$$MSE_{ki} = \frac{1}{T}\sum_{t}\left(a_k - f_{ki}\right)^2, \tag{6}$$

where a_k denotes actual earnings of firm k, and f_{ki} denotes the predicted earnings of firm k by model i. However, if forecast errors are measured in terms of levels of earnings, as the level of earnings increases in absolute magnitude, so will the absolute magnitude of the forecast errors. In addition, Dacco and Satchell (1999) argue that MSE measure may be not quite appropriate for the non-linear models since this measure may imply that a non-linear model is less accurate than a linear one when it is not actually true. Accordingly, I calculate a second scale invariant measure of accuracy – MSPE:

$$MSPE_{ki} = \frac{1}{T}\sum_{t}\left(\frac{a_k - f_{ki}}{a_k}\right)^2. \tag{7}$$

In order to compare the MSE and MSPE error measures from different models, I use the asymptotic loss differential test proposed by Diebold and Mariano (1995). The test considers a sample path $\{d_t\}_{t=1}^{T}$ of a loss-differential series and tests the null

[4] I use two companies to find the most appropriate specifications of neural networks for this study. These companies are not from the sample of forty-eight companies considered in the results.

hypothesis of equal forecast accuracy between two alternative models by exercising the next statistic:

$$S = \frac{\bar{d}}{\sqrt{2\pi\hat{f}_d(0)/T}} \sim N(0,1), \tag{8}$$

where $\bar{d}$ is the sample mean loss differential:

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T}\left[g(y_t, \hat{y}_{it}) - g(y_t, \hat{y}_{jt})\right], \tag{9}$$

and $\hat{f}_d(0)$ is a consistent estimator of the spectral density at frequency *0*. It is computed as a weighted sum of the available sample autocovariences:

$$\hat{f}_d(0) = \frac{1}{2\pi}\sum_{\tau=-(T-1)}^{(T-1)} I\left(\frac{\tau}{S(T)}\right)\frac{1}{T}\sum_{t=|\tau|+1}^{T}\left(d_t - \bar{d}\right)\left(d_{t-|\tau|} - \bar{d}\right), \tag{10}$$

where the uniform lag window *I( )* is given by:

$$I\left(\frac{\tau}{S(T)}\right) = 1 \quad for \quad \left|\frac{\tau}{S(T)}\right| \le 1 \tag{11}$$
$$= 0 \quad otherwise.$$

The truncation lag, *S(T)*, is equal to zero for one-year-ahead forecasts and equal to one for two-year-ahead-forecasts. It follows from the familiar fact that *k*-step-ahead forecast errors are at most *(k-1)* dependent. I define the loss differential series to be $d_t = (a_t - f_{it}) - (a_t - f_{jt})$ for the MSE test and $d_t = (1 - f_{it}/a_t) - (1 - f_{jt}/a_t)$ for the MSPE test, where a_t denotes actual earnings at time *t*, while f_{it} and f_{jt} are predicted earnings by models *i* and *j,* respectively. The formula indicates that due to the cumulation of autocovarience terms, the correction for serial correlation may be substantial even if the loss differential is only weakly correlated.

IV. Empirical Results

A. *Comparison of forecast accuracy.*

Looking at earnings patterns of different companies between 1973 and 1992, it seems natural to divide them into two subgroups, according to their earnings volatility. To describe the earnings volatility quantitatively, I employ the coefficient of variation of the first difference of earnings:

$$CV_k = \frac{Q_3(d_k(a)) - Q_1(d_k(a))}{Q_2(d_k(a))}, \tag{12}$$

where $Q_2(d_k(a))$ denotes the median change in earnings of firm *k*, and $Q_3()$ - $Q_1()$ is the interquartile range of these changes. The greater the coefficient, the more volatile changes in earnings, and therefore, the harder the task of forecasting. I compute the coefficient of variation for each firm and divide companies into two groups consisting of firms with values of the coefficient above and below its median value. I call these groups the *"cycle"* and the *"growth"* groups, respectively. Each group contains twenty-four companies. For purposes of illustration, I present earnings of the representative cycle and growth group firms in Figure 1. While earnings are highly cyclical and volatile in the first graph, earnings, in the second graph, grow remarkably in line with S&P500 earnings. Thus, the growth group consists of firms whose earnings move together with the market, whereas earnings of the cycle group firms are more susceptible to economic fluctuations.

First, using forecasting models described above, I generate one- and two-year-ahead forecasts of earnings per share for the 1993-2002 period. Next, I calculate MSE and MSPE for each company, forecast horizon and forecasting method. Then, I carry out the Diebold and Mariano test of equal forecasting accuracy and assign ranks from one to five to each of the five models in consideration[5]. Finally, I sum the ranks of each forecasting method i across N firms:

$$rank_i = 1/N \sum_{k=1}^{N} rank_{ki} \ . \qquad (13)$$

I present the results in Table 1 and Table 2. The smaller the rank of a model, the better predictive accuracy the model obtains. Note that "average across models" ranks are not equal to three, since it is often the case that by performing the Diebold and Mariano test, I cannot reject the hypothesis of equal forecasting accuracy.

The main result in Table 1 is that artificial neural networks that make use of quarterly data (QNN) outperform the other time-series methods in consideration and produce the comparable performance to financial analysts. Average ranks of QNN based on the MSE and MSPE error measures are 1.40 and 1.47, while financial analysts' ranks are 1.69 and 1.91, respectively. Financial analysts generate the better accuracy than the random walk with drift model (RW) based on the MSE measure, 1.69 versus 2.23, and comparable accuracy based on the MSPE measure, 1.91 versus 2.03. On the other hand, artificial neural networks that exploit annual data (ANN) and their linear analog, the Box and Jenkins procedure (BJ), fail to produce a better accuracy than the random walk. I link the poor performance of ANN relative to the QNN model to the insufficient number of data points used for its training. In contrast, I explain the success of QNN by two factors: the desegregation effect that results from higher data frequency and the ability to avoid significant outlier quarters in training without radically reducing the training set. The linear BJ procedure produces the worst accuracy results.

Table 2 presents the results for the two-year-ahead forecast horizon. The QNN modeling technique continues to be a leader in forecasting accuracy based on both error measures. Its ranks are the smallest and equal to 1.73 and 1.77 for the MSE and MSPE error measures. Taking into account that in order to produce a two-year-ahead forecast, we need to obtain one- to eight-quarters-ahead forecasts and the fact that the forecast accuracy decreases with the forecast horizon, the QNN superior accuracy is a prominent

[5] Note that if I cannot reject the hypothesis of equal forecasting accuracy between models i and j at the *5%* significance level, both models get the same rank.

result. On the contrary, financial analysts produce worse accuracy two-year-ahead forecasts made at the beginning of a fiscal year than any other model in consideration; even the linear BJ procedure and the adaptive random walk supply more accurate forecasts. This fact sheds significant doubt on the credibility of financial analysts as providers of accurate long-term earnings forecasts.

By looking at subgroups, we can draw the following conclusions. For the one-year-ahead forecast horizon, analysts' forecasts have the accuracy comparable to the random walk forecast accuracy for the cycle group companies and have better accuracy for the growth group companies. For the two-year-ahead forecast horizon, the analysts' forecast accuracy is inferior to all models for the cycle group companies and only comparable to the BJ model accuracy for the growth group companies. Next, for both forecast horizons, the advantage of quarterly neural networks is greater for the cycle group companies. According to both error measures, a gap between the QNN and the analysts' accuracy widens as we move from the growth to the cycle companies group. The average rank of the QNN model for the cycle group is about 36% (49%) smaller for the one- (two)-year-ahead forecast horizon than that of financial analysts. This is a focal result. It shows that neural networks are the most valuable in forecasting earnings of the high change in earnings volatility companies. It is apparently caused by their superior ability to extract nonlinear systematic patterns from series of past earnings. On the other hand, financial analysts consider larger sets of information that often consist of contradictory signals about companies' future prospects. As a result, they attach reduced weights to the information in histories of earnings and underestimate its importance. With respect to the growth group companies, the performance of neural networks is significantly undermined by sharp changes in earnings generating processes in the beginning of the 1990s, when relatively flat earnings plateaus were replaced by steady growth.

Finally, note that I perform an *ex ante* division of companies between the cycle and the growth groups. The coefficient of variation (12) is evaluated only using the 1973-1992 earning data. To verify the results using an *ex post* measure of variation, I reclassify companies between groups according to the coefficient of variation that is estimated using the 1983-2002 earnings data. In this case, the quarterly neural networks accuracy advantage is even more prominent for the cycle group firms as compared to the growth group companies.

To summarize, according to the rank orders forecast comparison procedure, neural networks utilizing quarterly data appear to be the method with the best accuracy of one- and two-year-ahead EPS forecasts made at the beginning of a fiscal year. Their advantage is the most evident for the high change in earnings volatility companies whose earnings regularly deviate from the market. On the contrary, financial analysts predict relatively well earnings of companies whose earnings move in line with market earnings. This result suggests that financial analysts generally predict the overall market component, but have a lack of ability to foresee specific fluctuations. Do they fail to predict the upward or downward deviations, or both? The next section provides an insight into this interesting question.

Model	Average Rank based on MSE			Average Rank based on MSPE		
	“Cycle” firms	“Growth” firms	Total	“Cycle” firms	“Growth” firms	Total
Financial Analysts	2.08	1.29	1.69	2.12	1.70	1.91
RW with drift	*2.08*	*2.38*	*2.23*	*1.93*	*2.13*	*2.03*
QNN (Quarterly EPS)	1.29	1.50	1.40	1.38	1.55	1.47
ANN (Annual EPS)	2.33	2.08	2.20	2.04	2.60	2.32
BJ	2.29	2.92	2.60	1.98	2.66	2.32
Average across models	2.01	2.03	2.02	1.89	2.13	2.01

Table 1. Rank orders of financial analysts and time-series models forecasting *one-year-ahead EPS* at the beginning of a fiscal year. Forecast accuracy is measured by MSE and MSPE. The Diebold-Mariano predictive accuracy test is applied to MSE and MSPE loss differentials. The ranks are assigned according to the *5%* significance level.

Model	Average Rank based on MSE			Average Rank based on MSPE		
	“Cycle” firms	“Growth” firms	Total	“Cycle” firms	“Growth” firms	Total
Financial Analysts	3.33	2.63	2.98	3.29	2.75	3.02
RW with drift	*2.21*	*2.58*	*2.40*	*2.21*	*2.33*	*2.27*
QNN (Quarterly EPS)	1.63	1.83	1.73	1.75	1.79	1.77
ANN (Annual EPS)	2.17	2.25	2.21	1.92	2.34	2.13
BJ	1.96	3.04	2.50	1.88	2.54	2.21
Average across models	2.26	2.47	2.37	2.21	2.35	2.28

Table 2. Rank orders of financial analysts and time-series models forecasting *two-year-ahead EPS* at the beginning of a fiscal year. Forecast accuracy is measured by MSE and MSPE. The Diebold-Mariano predictive accuracy test is applied to MSE and MSPE loss differentials. The ranks are assigned according to the *5%* significance level.

Model	One-year-ahead			Two-year-ahead		
	Prediction matrixes Up	Prediction matrixes down	HM p-value	Prediction matrixes up	Prediction matrixes Down	HM p-value
Financial Analysts	318, 12	110, 40	0.00	295, 5	116, 16	0.00
RW with drift	*267, 63*	*117, 33*	*0.20*	*244, 56*	*110, 22*	*0.06*
QNN (Quarterly EPS)	228, 102	54, 96	0.00	224, 76	52, 80	0.00
ANN (Annual EPS)	238, 92	69, 81	0.00	223, 77	61, 71	0.00
BJ	228, 102	107, 43	0.82	238, 62	95, 37	0.06
Total	330	150		300	132	

Table 3. Prediction matrixes of financial analysts and time-series models forecasting *one-* and *two-year-ahead EPS* at the beginning of a fiscal year. The first entry corresponds to correctly predicted up moves, second to actual up/predicted down, third to actual down/predicted up and fourth to correctly predicted down moves. *HM p-values* for the rejection of the hypothesis of *no forecasting skills*.

B. *The direction of change measure.*

A slightly different approach to assess the forecast accuracy and to get an insight into sources driving statistical models and financial analysts' forecasting abilities is to utilize the direction of change measure. This measure is related to forecasts interpreted only in terms of whether a firm's earnings will increase or decrease. I demonstrate the performances of models in terms of prediction matrixes in Table 3. They portray forecasts as the numbers of correct and incorrect predictions of the direction of change[6].

As it is evident from the results, financial analysts produce fewer mistakes in predicting upward movements (actual up/predicted down is equal to twelve and five for the one- and two-year-ahead forecast horizon), but more mistakes in predicting downward movements (actual down/predicted up is equal to 110 and 116 for the one- and two-year-ahead forecast horizon) as compared to neural networks. In fact, analysts correctly predict 96% (98%) of one- (two)-year-ahead up moves and only 27% (12%) of one- (two)-year-ahead down moves, whereas similar statistics for quarterly neural networks are 69% (75%) and 64% (61%). Quarterly neural networks possess the best skills for predicting down moves (96 out of 150 and 80 out of 132 for the one- and two-year-ahead forecast horizon, respectively). It apparently leads to their superior performance observed in terms of the rank orders. On the contrary, financial analysts often miss the correct prediction of downward movements, which are the most important deviations to predict in the environment of rising earnings. Financial analysts have a tendency to produce upward predictions and, thus, to some extent, ignore histories of earnings. In the case of the cycle group firms, these histories may contain a number of long lasting downturns pointing at a great potential for downward deviations from the overall market in the future.

Next, note that there are some similarities between financial analysts and the linear BJ and random walk models. Namely, in the case of one-year ahead forecast horizon, the number of correctly predicted down moves by analysts is only slightly higher than that by the random walk, 40 versus 33 out of 150, and similar to the BJ model predictions, 40 and 43, respectively. For the two-year-ahead forecast horizon, analysts correctly predict 16 out of 132 down moves, while the random walk model correctly predicts 22. This suggests that financial analysts incorporate trends into their forecasts that shadow drift components of the random walk model. This behavior is similar to the behavior of the naive investor, who extrapolates the past performance into the future[7].

Finally, there is the question of whether the least confusing models are the models that we would choose based on the MSE and MSPE forecast measures in the setting of real-time forecasting. To provide an answer, I perform the nonparametric test given by

[6] The first entry corresponds to correctly predicted up moves, second to actual up/predicted down, third to actual down/predicted up and fourth to correctly predicted down moves.

[7] See, for example, Lakonishok, Shleifer and Vishny (1994). They argue that value strategies yield higher returns because these strategies exploit the suboptimal behavior of the typical investor, who extrapolates past earnings growth too far into the future.

Hendrickson and Merton (1981) and compute HM p-values for the rejection of hypothesis of no forecasting skills. According to the results, I reject the hypothesis of no forecasting skills for financial analysts and neural networks models for both forecast horizons at the *1%* significance level. These models are found to be useful as predictors of the sign of change in earnings. On the contrary, the random walk model, which demonstrates a solid performance based on the MSE and MSPE measures, has no forecasting skills for the direction of change.

C. *Relative informational content of forecasts.*

The adaptive statistical models used to forecast earnings exploit only series of past earnings, whereas financial analysts make use of a considerably broader information set. It constitutes the information advantage. Then, how should we interpret the differences in forecasts? Does each model have strength of its own, or do financial analysts' forecasts dominate in the sense of incorporating all information in the other model plus sum? I examine this question by considering the regression of actual changes in earnings on changes forecasted by financial analysts and statistical models. This procedure may comprise advantages over the direct comparison of MSE or MSPE error measures. For example, if the MSE are close for two forecasts, and performing the Diebold and Mariano test, we cannot reject the hypothesis of equal forecasting accuracy; little can be concluded about the relative merits of the two. Furthermore, even if the MSE of one model is bigger than the other, it may still be the case that its forecasts contain additional information. There is no way to test for this using the MSE framework. Therefore, I consider the following regression equation:

$$a_{kt} - a_{k(t-s)} = \alpha + \beta_1 (f_{kt}^1 - a_{k(t-s)}) + \beta_2 (f_{kt}^2 - a_{k(t-s)}) + \varepsilon_{kt} \qquad (14)$$

where a_k denotes actual earnings of firm k, f_k^1 and f_k^2 are predicted earnings by models one and two, while $s = 1, 2$ is the forecast horizon. If neither model contains useful information for s-period-ahead forecasts, then estimates of β_1 and β_2 should both be zero, and α would be the average s-period-ahead change in earnings. If forecasts are not perfectly correlated, and both models contain independent information, then β_1 and β_2 should both be nonzero. Finally, if the model two is completely contained in the model one, and the model one contains further relevant information as well, then β_2 but not β_1 should be nonzero.

I focus on the performance of financial analysts (model 1) versus the quarterly neural networks (model 2), which were shown to have superior predictive abilities in terms of rank orders, and the random walk model, which represents a sufficiently simple forecasting technique. The procedure consists of estimating Equation (14) first, and then, testing the hypotheses: $H0: \beta_1 = 0$ that analysts' forecasts contain no information, which is not incorporated in a constant term and in statistical models forecasts, and $H0: \beta_2 = 0$ that the statistical models contain no information, which is not included in a constant term and in analysts' forecasts. Note that it does not seem reasonable to estimate

Equation (14) for each company or each year separately[8]. Therefore, I consider the pooled data set that produces 480 (432) data points in the case of one- (two)-year-ahead forecast horizon. The OLS estimator is not the best linear unbiased estimator in this case. Therefore, I consider the feasible generalized least square estimator, which is the weighted average of between- and within-group estimators given in Maddala (1971):

$$\hat{\beta}_{GLS} = \Delta\hat{\beta}_b + (I_k - \Delta)\hat{\beta}_w \,, \tag{15}$$

where consistent estimators of unknown σ_θ^2 and σ_v^2 are used to determine the weight *Δ*. This is called the random effect model, where $\varepsilon_{it} = \theta_i + v_{it}$, and v_{it} are treated as random variables.

Table 4 presents the estimated coefficients of Equation (14) for the one- and two-year-ahead forecast horizon. With respect to the one-year-ahead forecast horizon, the coefficient estimates for financial analysts' and RW/QNN forecasts are both nonzero and statistically significant for the cycle group companies. It indicates that there is some information in RW/QNN forecasts that is not in analysts' forecasts. This result suggests that financial analysts process the information in histories of earnings differently than the time-series models or that they neglect this information at all. It contradicts the common view that analysts make use of all available information in constructing their forecasts.

Now, consider two-year-ahead forecasts. For the cycle group firms, random walk model has independent information; its coefficient estimate is nonzero and statistically significant. Moreover, while the coefficient estimate for quarterly neural networks forecasts is nonzero and statistically significant, the coefficient estimate for financial analysts' forecasts is not significantly different from zero at the *5%* level. It reveals that analysts' forecasts contain no information, which is not incorporated in a constant term and in QNN forecasts, for companies with volatile earnings. This result demonstrates that quarterly neural networks forecasts are not collinear with financial analysts' forecasts and that the difference between the QNN and financial analysts' accuracy is meaningful.

To summarize, quarterly neural networks as well as the adaptive random walk with drift model contain information not in a constant term and in analysts' forecasts for the high change in earnings volatility companies. In contrast, neither coefficient estimate for the *Time-Series* variable in Table 4 is significant for the low change in earnings volatility companies indicating that all information is already included in financial analysts' forecasts. Overall, without considering conventional measures of accuracy, results support the hypothesis that financial analysts have relatively good predictive abilities for companies with steady earnings that follow the market, whereas quarterly neural networks carry useful information that is not in financial analysts' forecasts for companies with volatile earnings. It poses a challenge to the previously acclaimed notion of financial analysts' informational superiority.

Next, to get an insight into the properties of financial analysts' forecasts, I compute the forecast bias for financial analysts and RW/QNN forecasting methods. I employ a modification of Equation (14) and regress the predicted minus the actual

[8] The sample consists of ten forecast dates and twenty-four companies in each group.

change in earnings on a constant. Table 5 presents the results. With respect to the growth group companies, financial analysts produce the positive but not statistically significant bias for the one-year-ahead forecast horizon and the positive and significant bias for the two-year-ahead forecast horizon. The random walk and quarterly neural networks models have negative and in two instances statistically significant biases, which are apparently caused by the relatively flat earnings plateau in the beginning of the 1990s and the consecutive shift to the drastic growth as demonstrated in Figure 1. For the cycle group firms, financial analysts generate large optimistic biases: *27* cents for the one-year-ahead and *53* cents for the two-year-ahead forecast horizon. These estimates are significantly different from zero and more than three times larger than those for the growth group firms are. It implies that financial analysts tend to predict the upward growth for all types of companies, which results in the large optimistic bias for the cycle group firms. On the contrary, quarterly neural networks and the random walk model are not biased. The estimates of the forecast bias are not statistically significant for both forecast horizons. These facts help to explain the observed neural networks accuracy advantage for the volatile change in earnings firms. Overall, the results suggest that the analysts' forecast bias increases with the volatility of earnings. To my knowledge, this observation is a new result in the literature examining the analysts' forecast rationality. It deserves further attention in a separate study.

D. *Number of analysts and the standard deviation of individual forecasts.*

While some companies may be followed by over than thirty financial analysts, others are covered by only few. For some companies, the standard deviation of individual analysts' forecasts may be small indicating that the majority of analysts agree on estimates, for others, it may be considerably large pointing at disagreement between financial analysts and uncertainty about companies' prospects. Note that even though the number of analysts issuing forecasts and the standard deviation of individual forecasts for a specific company may change from year to year, it is very persistent.

One would naturally expect consensus forecasts to have better accuracy relative to time-series models as the number of individual analysts' forecasts used to construct consensus increases. This view can be explained by widening the information set with each additional analyst or by portfolio benefits of averaging. Similarly, the more disperse individual analysts' forecasts used to construct consensus forecasts, the more likely there is to be substantial uncertainty about future earnings. It may be a sign that different analysts receive different signals about future prospects, or that they process the information flow differently. As a result, we can anticipate the relative accuracy of financial analysts' consensus forecasts to increase as more analysts follow the company and as the standard deviation of individual forecasts decreases.

To test these ideas, I calculate the mean number of forecasts and the mean standard deviation of forecasts for each company and forecast horizon over a ten- (nine)-year period. The number of analysts who issue one- (two)-year-ahead forecasts in the sample is as many as thirty-four (twenty-one), and as few as five (two). Therefore, to

Time-Series Model	One-year-ahead					
	"Cycle" firms			"Growth" firms		
	Const	FA	Time-Series	Const	FA	Time-Series
RW with drift	-0.14	0.78	-0.93	-0.05	0.80	0.35
	(3.15)	(10.3)	(3.30)	(1.35)	(9.91)	(1.76)
QNN	-0.17	0.69	0.57	-0.01	0.75	0.13
	(4.25)	(9.37)	(5.90)	(0.43)	(8.67)	(1.53)
	Two-year-ahead					
RW with drift	-0.10	0.43	-0.63	0.04	0.46	0.15
	(1.09)	(4.04)	(2.75)	(0.61)	(4.38)	(1.41)
QNN	-0.10	0.23	0.53	0.04	0.47	0.21
	(1.13)	(1.88)	(5.26)	(0.58)	(3.95)	(1.87)

Table 4. Informational content of forecasts. Financial analysts versus the adaptive time-series models: Estimation of Equation (14). FA is financial analysts' predictions minus actually realized values. Time-Series is random walk with drift/quarterly neural networks predictions minus actually realized values. Heteroskedasticity-robust *t*-statistics in absolute value are in parentheses.

Model	One-year-ahead		Two-year-ahead	
	"Cycle" firms	"Growth" firms	"Cycle" firms	"Growth" firms
	Const	Const	Const	Const
Fin. Analysts	0.27	0.05	0.53	0.15
	(4.07)	(1.74)	(4.51)	(2.86)
RW with drift	0.02	-0.07	0.05	-0.06
	(0.31)	(2.15)	(0.64)	(1.53)
QNN	-0.05	-0.06	-0.03	-0.12
	(0.75)	(1.70)	(0.29)	(2.18)

Table 5. Forecast bias for financial analysts, random walk with drift and quarterly neural networks *one-* and *two-year-ahead EPS* forecasts made at the beginning of a fiscal year. The bias is estimated by the regression of the predicted minus the actual change in earnings on a constant; *t*-statistics in absolute value are in parentheses.

Model	One-year-ahead			Two-year-ahead		
	Const	$Number_{FA}$	SD_{FA}	Const	$Number_{FA}$	SD_{FA}
FA – RW	-0.12	0.01		0.04	0.00	
	(1.00)	(0.53)		(0.35)	(0.55)	
	0.04		0.70	-0.12		0.85
	(1.01)		(5.12)	(1.38)		(1.56)
FA – QNN	-0.08	0.01		0.01	0.01	
	(1.55)	(0.79)		(0.09)	(0.58)	
	-0.18		1.26	-0.18		2.04
	(2.40)		(1.86)	(1.89)		(3.50)

Table 6. Relative forecast accuracy and the number of analysts issuing forecasts, standard deviation of individual forecasts used to construct consensus forecasts. *FA-RW, FA-QNN* are the differences in the RMSE of financial analysts and the random walk with drift/quarterly neural networks models scaled by the standard deviation of the change in earnings over the 1993-2002. $Number_{FA}$ is the mean number of financial analysts issuing forecasts, SD_{FA} is the mean standard deviation of individual forecasts over the 1993-2002. Heteroskedasticity-robust *t*-statistics in absolute value are in parentheses.

control for possible problems with the standard deviation as a proxy for uncertainty when only few analysts follow the company, I exclude two companies for which the mean number of analysts is less than three.

The next question I need to address is which measure of the relative accuracy to employ. While subtracting the MSPE error measures of two alternative models can produce meaningless values, subtracting the MSE measures does not look practical as well[9]. Therefore, I suggest the following measure of relative forecast accuracy. First, I take differences of RMSE calculated over the 1993-2002 period and then, scale it by the standard deviation of the first difference in earnings over the same period, $(RMSE_{Fin.An.} - RMSE_{Time-Series})/\sigma(d(A))$. The idea is that differences in RMSE should be discounted more as earnings are less predictable. The proposed measure should allow us to reveal the effects of the number of analysts and the standard deviation of individual forecasts on the relative forecast accuracy without distorting the results by the difficulty of the forecasting task. Finally, I exclude outlier observations and regress the constructed earlier measure of the relative forecast accuracy on a constant and the number of analysts or the standard deviation of individual forecasts. Table 6 presents the results.

Conversely to intuition, the relative accuracy of financial analysts' consensus forecasts does not depend on the number of forecasts used to construct them. Instead of being negative as expected, the estimates of coefficients are not significantly different from zero for both time-series models and both forecast horizons. We do not observe benefits of aggregating a large number of individual analysts' forecasts on the relative performance of financial analysts. It implies that all analysts make use of identical sets of information in the constriction of forecasts. Furthermore, note that there is a positive relationship between the number of analysts issuing forecasts and the size of the company. The bigger the size of a company, the more information is available about the company, that is, the bigger the dimensionality of its information set. Therefore, the greater number of analysts issuing forecasts for a company may point to the situation when each additional analyst processes information available to other analysts plus some new information. In this case, the relative accuracy of analysts' forecasts would increase with the number of analysts. However, the observed results contradict this view.

With respect to the standard deviation of individual analysts' forecasts, as expected, the relative accuracy of financial analysts tends to increase as the dispersion of individual forecasts decreases. The estimates of coefficients are always positive and statistically significant for the random walk model in the case of one-year-ahead forecast horizon and for quarterly neural networks in the case of the two-year-ahead forecast horizon. It indicates that the value of adaptive time-series models forecasts increases with the standard deviations of individual analysts' forecasts and, as a result, with the degree of uncertainty about future prospects. Note that the improvement in the relative accuracy of adaptive time-series models can be linked not only to the analysts' disagreement, but

[9] In the case of the MSPE error measure: if the denominator value is close to zero, one of the MSPEs can be quite large. In the case of the MSE error measure: similar resulting differences in errors are not the same for companies with different volatilities of earnings.

to the volatility of earnings as well. Indeed, as the standard deviation of the change in earnings increases, the standard deviation of individual analysts' forecasts grows too. The coefficient of correlation between the two is high and statistically significant. Thus, quarterly neural networks perform relatively better for companies with the high dispersion of individual analysts' forecasts, that is, for companies with highly volatile earnings. Furthermore, the estimate of the coefficient (2.04) is greater for the two-year-ahead forecast horizon than that (1.26) for the one-year-ahead forecast horizon indicating that the neural networks relative accuracy increases faster for longer forecast horizons. These results are absolutely in line with the prior findings.

To conclude, I find that the relative accuracy of financial analysts is not related to the dimensionality of the information set as measured by the mean number of analysts issuing forecasts, but rather positively related to the quality of this set as measured by the mean standard deviation of individual analysts' forecasts. The results suggest that not the size, but the type of the company is a main determinant of the financial analysts' relative forecast accuracy.

V. Conclusions

According to the rank orders procedure that utilizes the MSE and MSPE error measures, neural networks exploiting quarterly data produce forecasts that are comparable in accuracy to analysts' predictions in the case of one-year-ahead forecast horizon or even superior in the case of two-year-ahead forecast horizon. In fact, financial analysts produce less accurate two-year-ahead forecasts made at the beginning of a fiscal year than any other adaptive time-series model in consideration; even the linear random walk with drift model supplies more accurate forecasts. This fact sheds significant doubt on the credibility of financial analysts as providers of accurate long-term earnings forecasts at the beginning of a fiscal year.

Furthermore, I show that the accuracy advantage of artificial neural networks is more pronounced for companies with volatile and, therefore, hardly predictable changes in earnings. In contrast, financial analysts produce forecasts of a relatively good accuracy for companies with steady earnings, which generally move in line with market earnings. It suggests that analysts primarily predict the overall market component, but often fail to foresee firm specific fluctuations. Overall, the results imply that financial analysts have either inferior abilities relative to artificial neural networks to extract nonlinear systematic patterns from histories of volatile earnings or a lack of incentives to extract it properly, or both.

As is evident from the direction of change measure, both financial analysts and neural networks are found to be useful predictors of the sign of change in earnings. However, the structure of predictions differs. Financial analysts have the advantage in recognizing upward movements, whereas quarterly neural networks possess the best skills to recognize downward moves. The financial analysts' tendency to produce mostly positive sign predictions is consistent with generally growing earnings in the 1990s, but hides a potential danger to ignore histories of earnings, which contain frequent downturns for the cycle group companies. On the other hand, the better ability of quarterly neural networks to predict downward movements may lead to their superior performance observed in terms of rank orders.

Next, without considering conventional measures of accuracy, I present evidence suggesting that adaptive statistical models that only utilize series of past earnings contain information not in a constant term and in analysts' forecasts made at the beginning of a fiscal year for companies with volatile earnings. It indicates that analysts overlook the information in histories of earnings. The observed behavior may be caused by analysts being abundant with larger sets of noisy information. Therefore, they assign smaller weight to the information in series of past earnings. This result contradicts the widely accepted idea that financial analysts use all available information in constructing forecasts and poses a challenge to the previously acclaimed notion of financial analysts' informational superiority.

The last question I study is how the number of analysts issuing forecasts and the standard deviation of individual forecasts influence the relative accuracy of financial analysts' consensus forecasts. I find that the relative accuracy of financial analysts depends on the type of the company, as measured by the volatility of its earnings, and not the size of the company. Namely, the relative accuracy of analysts' consensus forecasts is negatively related to the standard deviation of individual forecasts and independent of the number of forecasts used to construct consensus. It reveals that the amount of available information is not a main determinant of the relative forecast accuracy.

Finally, I want to conclude with a warning about the interpretation of the results. The fact that one forecasting technique does well or poorly for one sample period does not necessarily mean that it will do well or poorly in the future periods. The results could change if the structure of the economy is changing, which is of course true of any econometric result.

References

Albrecht, W., Lookabill, L., and McKeown, J., 1977, The time-series properties of annual earnings, *Journal of Accounting Research*, 15, 226-44.

Ball, R., and Watts, R., 1972, Some time-series properties of accounting income, *Journal of Finance*, 27, 663-81.

Barron, A., 1991, Universal approximation bounds for superpositions of sigmoidal function, Report No. 58, Department of Statistics, University of Illinois, Urbana-Champaign.

Bishop, C., 1995, Neural Networks for Pattern Recognition, Oxford: University Press.

Box, G., and Jenkins, G., 1970, Time Series Analysis, Forecasting and Control, San Francisco: Holden-Day, Inc.

Brown, L., and Roseff, M., 1978, The superiority of analysts forecasts as measures of expectations: evidence from earnings, *Journal of Finance*, 33, 1-16.

Brown, L., and Roseff, M., 1979, Univariate time series models of quarterly accounting earnings per share: a proposed model, *Journal of Accounting Research*, 17, 178-189.

Brown, L., Richardson, G., and Schwager, G., 1987, An information interpretation of financial analysts superiority in forecasting earnings, *Journal of Accounting Research*, 25, 49-67.

Collins, W., and Hopwood, S., 1980, A multivariate analysis of annual earnings forecasts generated from quarterly forecasts of financial analysis and univariate time series models, *Journal of Accounting Research*, 1980, 18, 390-406.

Cybenko, G., 1989, Approximation by superpositions of a sigmoidal function, *Mathematics of Control, Signals, and Systems*, 2, 303-314.

Dacco, R., and Satchell, S., 1990, Why do regime-switching models forecast so badly, *Journal of Forecasting*, 18, 1-16.

Diebold, F., and Mariano, R., 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.

Dugar, A., and Nathan, S., 1995, The effects of investment banking relationships on financial analysts' earnings forecasts and investment recommendations, *Contemporary Accounting Research*, 12, 131-660.

Fausett, L., 1994, Fundamentals of Neural Networks, New York: Prentice Hall.

Francis, J., and Philbrick, D., 1993, Analysts' decisions as products of a multi-task environment, *Journal of Accounting Research*, 31, 216-230.

Fried, D., and Givoly, D., 1982, Financial analysts' forecasts of earnings: a better surrogate for earnings expectations, *Journal of Accounting and Economics*, 4, 85-107.

Henriksson, R., and Merton, R., 1981, On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills, *The Journal of Business*, 54, 513-533.

Hong, H., Kubik, J., and Solomon, A., 2000, Security analysts' career concerns and herding of earnings forecasts, *RAND Journal of Economics*, 31, 121-144.

Hopwood, W., McKeown, J., and Newbold, P., 1982, The additional information content of quarterly earnings reports: intertemporal disaggregation, *Journal of Accounting Research*, 20, 343-349.

Hornik, K., Stinchcombe, S., and White, H., 1989, Multilayer feedworward networks are universal approximators, *Neural Networks*, 2, 359-366.

Kuan, C., and Liu, T., 1995, Forecasting exchange rates using feedforward and recurrent neural networks, *Journal of Applied Econometrics*, 10, 347-364.

Lakonishok, J., Shleifer, A., and Vishny, R., 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance*, 49, 1541-1578.

Lin, H., and McNichols, M., 1998, Underwriting relationships, analysts' earnings forecasts and investment recommendations, *Journal of Accounting and Economics*, 25, 101-127.

Maddala, G., 1971, The likelihood approach to pooling cross-section and time series data, *Econometrica*, 39, 939-953.

McNichols, M., and O'Brian, P., 1997, Self selection and analyst coverage, *Journal of Accounting Research*, 35, 167-199.

Rumelhart, D., and McClelland, J., 1986, Parallel Distributed Processing, Vol. 1, Cambridge, MA: MIT Press.

Swanson, N., and White, H., 1995, A model selection approach to assessing the information in the term structure using linear models and artificial neural networks, *Journal of Business and Economic Statistics*, 13, 265-275.

Swanson, N., and White, H., 1997, A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks, *The Review of Economics and Statistics*, 79, 540-550.

Tkacz, G., 2001, Neural network forecasting of Canadian GDP growth, *International Journal of Forecasting*, 17, 57-69.

Watts, R., and Leftwich, R., 1977, The time series of annual accounting earnings, *Journal of Accounting Research*, 15, 253-71.