INDUSTRY BASED FUNDAMENTAL ANALYSIS: USING NEURAL NETWORKS AND A DUAL-LAYERED GENETIC ALGORITHM APPROACH.

By

Charles T. Stivason

Dissertation submitted to the faculty of

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Business Administration

with a major in Accounting and Information Systems

Dr. Tarun K. Sen, Chairman

Dr. John J. Maher

Dr. Robert T. Sumichrast

Dr. Robert M. Brown

Dr. Cintia M. Easterwood

INDUSTRY BASED FUNDAMENTAL ANALYSIS: USING NEURAL NETWORKS AND A DUAL-LAYERED GENETIC ALGORITHM APPROACH.

(ABSTRACT)

This research tests the ability of artificial learning methodologies to map market returns better than logistic regression. The learning methodologies used are neural networks and dual-layered genetic algorithms. These methodologies are used to develop a trading strategy to generate excess returns. The excess returns are compared to test the trading strategy's effectiveness. Market-adjusted and size-adjusted excess returns are calculated.

Using a trading strategy based approach the logistic regression models generated greater returns than the neural network and dual-layered genetic algorithm models. It appears that the noise in the financial markets prevents the artificial learning methodologies from properly mapping the market returns. The results confirm the findings that fundamental analysis can be used to generate excess returns.

ACKNOWLEDGMENTS

I would like to acknowledge the financial support of the Accounting and Information Systems department. The insight and help provided by my committee and chairman, Tarun K. Sen.

I would like to thank my wife, Kelcey for her support during graduate school. I would like to thank my children, Bryce and Carson for their patience during this endeavor.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

INTRODUCTION

Motivation

Financial analysts have consistently used the accounting information reported in the financial statements to obtain insights about a firm's current performance and future prospects (Palepu, et. al., 1996). The financial statements are the most widely available data on public corporations' economic activities. Accounting information captures, in part, the historic economic events of business entities. According to Palepu, et. al. (1996), "a firm's financial statements summarize the economic consequences of its business activities." The accounting information provides insights into how well a firm's management responds to the available opportunities. These insights should help establish a firm's value. Establishing the proper relationship between a firm's value and the accounting information is a complex process that requires complex modeling techniques.

An important question to accountants is: how can a relationship be established between accounting information and a firm's value? One possible method is to develop a successful trading strategy from the accounting information. If accomplished, the usefulness of the current system of financial accounting is supported. Ou and Penman (1989) and Holthausen and Larcker (1992) developed trading strategies based on the prediction of the sign of earnings' changes and the sign of excess return measures using public data. Ou and Penman found an average market-adjusted excess return for the 1973-1983 period of 8.3 percent using the prediction of the sign of earnings' changes. Holthausen and Larcker found an average market-adjusted excess return for the 1978-1988 period of 4.3 percent using the prediction of the sign of excess return measures. This indicates accounting information is useful in developing a successful trading strategy.

Researchers have recently attempted to show that market returns follow a complex nonlinear process, best mapped by complex nonlinear models. Ou and Penman (1989) and Holthausen and Larcker (1992) used logistic regression to predict the sign of the earnings' change and the sign of the return. If complex nonlinear models can be developed to map market returns better than logistic regression, more profitable trading strategies can be developed. Researchers who have studied complexity in market returns include Scheinkman and LeBaron (1989), Freeman and Tse (1992), Corhay and Rad (1994), Zapranis and Refenes (1994) and Brorsen and Yang (1994).

Scheinkman and Lebaron (1989) examined several calendar anomalies. One anomaly is that prior weekly returns help predict future returns even though they are uncorrelated. The data in their study are consistent with a theory that some of the variation would come from nonlinearities as opposed to randomness. The data are inconsistent with the theory that proposes that the returns be generated by IID (independent and identically distributed) random variables. A strict “rational expectations” view indicates that the nonlinearities may well be the result of a “law”. A less strict view acknowledges that agents are constantly “learning about the "true" law.

Freeman and Tse (1992) address the question of whether a nonlinear model can explain the difference between the earnings' response coefficient and the price-earnings' coefficient better than the linear models. They assume that permanent earnings are more accurately forecast than transitory earnings, causing transitory earnings surprises to be concentrated in the tails of the UE (unexpected earnings) distribution thus showing that UR (unexpected returns) - UE is nonlinear. A linear model assumes that a firm's response to unexpected earnings is constant among the ranges of unexpected earnings. Freeman and Tse suggest that “the absolute value of unexpected earnings is negatively correlated with earnings persistence.” Accordingly, earnings surprises can be classified between permanent and transitory, and permanent unexpected earnings will have a greater impact than transitory changes. The results indicate that traditional linear models underestimate the value of permanent earnings surprises because “a linear model heavily weights high-magnitude, low-value transitory earnings at the expense of low-magnitude, high-value permanent earnings.” Freeman and Tse (1992) conclude that the nonlinear models have greater explanatory power.

Corhay and Rad (1994) indicate that prior research shows that distributions of stock returns are leptokurtic[1] and skewed in the US and London markets. They looked at the stock

[1] Kurtosis is the degree of peakedness of a distribution. Kurtosis is computed by taking the fourth moment of a distribution. A distribution with a high peak is called Leptokurtic.

indices in France, Germany, Italy, the Netherlands and the UK and determined that the skewness and kurtosis apply in these smaller markets as well. This is further evidence that daily returns exhibit nonlinear dependence that cannot be captured by the simple random walk model. Zapranis and Refenes (1994) use a tactical asset allocation problem to justify the need for nonlinear modeling. They test the ability of a neural network to generate a greater return than multiple linear regression by switching holdings between equities and cash and bonds and cash. The neural networks outperform the multiple linear regression both in and out-of-sample. Brorsen and Yang (1994) find that the daily stock index return's variances change over time, and have large heteroskedasticity. They develop models to remove the nonlinearity and leptokurtosis in the data but cannot remove all of the complexities, and cannot dismiss deterministic chaos as a model of daily stock returns.

Other researchers, such as Brock et al. (1991), Malliaris (1994), and Hsieh (1991) have suggested chaotic dynamics, a deterministic model that appears random, as a model of market returns. Statistical models, such as logistic regression, use predefined functional forms that are relatively simple compared to market complexities. Brock et al. (1991) review the prior literature and propose that a deterministic chaos may generate asset returns. Malliaris (1994) found that neural networks outperformed the random walk model in ten test sets of the Standard and Poor's index during 1989-90. This result supports the deterministic structure of the stock market in the test period. Brock et al. argue that this deterministic model is “chaotic dynamics.” Hsieh (1991) shows that stock returns are not IID as a whole or by decile portfolios. He also shows that the weekly Standard and Poor returns, the daily returns for 1982-89, and the 15-minute returns in 1988 are not IID. These findings support his argument that chaotic dynamics govern stock returns and that nonlinear models are best used to model these chaotic dynamics. These studies suggest that market complexities require complex modeling techniques.

One popular complex nonlinear model is a neural network. Neural networks use a network of nodes and connecting weights to represent the interaction between input and output parameters in a prediction model. The primary function of a neural network model is to assign appropriate weights to the input nodes of a network so that a weighted function of the input nodes predicts the outputs.

Neural networks are not constrained by predetermined functional form and can produce any functional form necessary to map complex phenomena (White, 1988; Dutta and Shekhar, 1988; Refenes, et al., 1994; Bansal et al., 1993; Malliaris, 1994; etc.). White (1988) made one of the first attempts to use neural networks in financial settings. He outlined an approach to use neural networks in an attempt to detect nonlinear regularities in asset prices. White looked at the daily returns for IBM. The R^2 results showed that the model outperformed a linear model in the training period. However, the neural network did not outperform the linear model using a holdout sample. White suggested that any comparisons be done with monetary returns and not R^2. The current study uses returns for comparison purposes.

Dutta and Shekhar (1988) used neural networks to predict corporate bond ratings. Neural networks consistently out performed the regression models in the complex setting of predicting bond ratings in both the training sample and holdout sample. The increase in performance indicates that the linear regression model is not as appropriate in as complex a setting such as bond ratings. Refenes et. al. (1994) developed a back propagation neural network model to predict the relative performance of a stock six months in advance. The neural network model outperformed the traditional statistical techniques for forecasting within the framework of the arbitrage pricing theory model for stock ranking. They conclude that this improvement is due to limitations of traditional techniques in applications with nonlinearities in the data set.

Bansal et. al. (1993) compared the performance of regression with neural networks for simulated data. They controlled the amount of errors in the data and found that the regressions had better performance in regard to R^2. However, the payoff of using a trading strategy is unaffected by changes in data quality for neural networks and outperforms the regression in terms of payoff. Their study did not include nonlinear relationships, which according to the authors explains why the regression models had superior performance using R^2. They state that nonlinear relationships in the data would cause the neural network R^2 performance to be better than the regression's performance.

Considering the complex nature of market returns and logistic regression's constraint of using a predefined functional form, neural networks should map the relationship between market returns and accounting information more accurately. Neural networks use additional layers to generate mappings that logistic regression does not. Instead of predicting the outcome directly from the input, neural networks add complexity through intermediate or hidden layers. These intermediate layers act as a filter to uncover patterns inherent in the sample data. Logistic regression estimates a partition of the sample space, which limits the ability of the model to map complex phenomena.

This complex structure of neural networks has associated problems. Neural networks lack explanatory capabilities, require a great amount of time to train properly, and provide precise mappings that may not generalize well unless the problem of over fitting is controlled. Another modeling technique that may capture the appropriate variables more precisely is the dual-layered genetic algorithm approach (Sikora and Shaw, 1994). This technique develops rules and strategies to explain the relationship between market returns and accounting information. This approach addresses neural networks' lack of explanatory capabilities (Eberhart, 1992) and produces a generalizable rule set. The general rule set should improve predictive accuracy over and above that provided by logistic regression. Richard Bauer (1994) suggests genetic algorithms as an alternative to neural networks in developing trading strategies. He uses genetic algorithms to successfully trade t-bonds and individual stocks. Murray Ruggiero (1997a) shows the profitability of using a trading strategy developed by genetic algorithms with t-bonds and the Swiss franc.

Both empirical and theoretical evidence suggest that complex nonlinear models are appropriate in market studies. This study examines the use of complex nonlinear models in trading strategies using accounting information. The traditional logistic regression is compared with neural networks and dual-layered genetic algorithms. The current study is primarily concerned with finding an appropriate form of modeling that is better than traditional statistical models to establish a relationship between public accounting information and returns.

Environmental factors may also impact the relative importance of accounting information. This study looks at the inclusion of environmental factors through an industry proxy. Prior research suggests that information varies in importance across industries (Downs, 1991; Biddle and Seow, 1991; Lustgarten and Thomadakis, 1980; Ely, 1991; Clarke, 1989). If information varies by industry then industry-specific models may be more appropriate than a market-wide model.

The rest of the dissertation is organized as follows. Chapter two reviews prior attempts to establish the appropriate relationship between accounting information and a firm's market value in order to generate abnormal returns. The most important aspect of the success of the model is the selection of the appropriate accounting information. Prior studies have used guided (Lev and Thiagarajan, 1993) and statistical searches to select accounting information for predicting excess returns (Ou and Penman, 1989; Holthausen and Larcker, 1992). The current study tests the ability of complex nonlinear modeling techniques to capture this relationship. Additionally, the study looks at the inclusion of environmental factors through an industry proxy.

Chapter three reviews the model development. The first section in chapter three addresses the variables included in the candidate list for the logistic regression model to predict excess returns. This study uses a "guided search" to select the independent variables. The binary dependent variable is the sign of abnormal returns for a prior twelve-month period. This variable is coded as one if the return was positive; zero otherwise. This is consistent with the Holthausen and Larcker (1992) study. The candidate variables are based on the empirical findings of Ou and Penman (O&P), Holthausen and Larcker (H&L) and Lev and Thiagarajan (L&T).

This study evaluates three modeling techniques to determine which best maps the complexity in financial markets. One complex modeling technique used is a neural network. Another technique used is a dual-layered symbolic approach using classification trees and genetic algorithms. The last technique used is the traditional logistic regression. This technique is used on a market-wide basis and an industry-specific basis. The independent and dependent variables used in these modeling techniques are discussed. The hypotheses regarding these techniques are presented along with the models used to test the hypotheses.

Chapter four reviews the complex modeling methodologies used. Back propagation, the most popular neural network algorithm, is described. This algorithm trains neural networks by presenting training data representing the phenomenon being learned. The chapter discusses how the algorithm adjusts weights between processing elements in the network to create a model. Next, the dual-layered genetic algorithm methodology is discussed. This approach combines the use of a recursive-partitioning algorithm with genetic algorithms. The recursive-partitioning algorithm develops a set of if-then rules used to classify all firms in the appropriate sample based on having positive or negative excess returns. This study uses the rule sets as the starting point for the genetic algorithms. Genetic algorithms use natural selection to solve optimization problems using training data. The three main components of genetic algorithms are discussed: 1) the coding scheme used to describe the problem in terms of a genetic code; 2) the way offspring are created through a crossover operator and mutation; and 3) the fitness function used to evaluate each offspring.

Chapter five reports the results of applying market-wide logistic regression models, industry-specific logistic regression models, neural network models, and dual-layered genetic algorithm models to develop a trading strategy to generate excess market returns. The conclusions, limitations and future research implications of the current study are discussed in chapter six.

CHAPTER 2
LITERATURE REVIEW

2.0 Introduction

The primary proposition of this study is that the relationship between accounting information and market value is best modeled with complex techniques. The specific modeling techniques used are neural networks and dual-layered genetic algorithms. Logistic regression is used as a base model. Trading strategies are developed from the model's ability to predict the directional change in returns for individual firms. The excess returns generated from using these trading strategies are used to measure the effectiveness of the modeling techniques.

This chapter is divided into five sections. The first section reviews literature that validates the possibility of mispricing. Financial analysts invest in security analysis to generate returns in excess of the associated costs of the analysis. This indicates that some mispricing exists, since these are knowledgeable agents undertaking this cost. The first section also reviews the use of accounting information and size in market studies. The studies included are related to "fundamental analysis." A key aspect of the models developed is the selection of the appropriate accounting information or variables. The articles reviewed discuss methods of variable selection. The primary studies of interest have used a guided search (Lev and Thiagarajan, 1993) and statistical searches (Ou and Penman, 1989; Holthausen and Larcker, 1992) to select accounting variables for predicting excess returns. The guided search justifies the selection of variables through analysts' discussions and prior research. The current study adopts this approach.

The second section discusses some market complexity problems. Market complexity problems have been established by a number of studies. One of the most significant problems that must be addressed in market studies is the nonlinearity of the market returns. Traditionally one of the easiest ways of establishing this significance is to compare linear models to nonlinear models. The literature (Scheinkman and LeBaron, 1989; Freeman and Tse, 1992; Corhay and Rad, 1994; Zapranis and Refenes, 1994; Brorsen and Yang, 1994; Brock et al., 1991; Hsieh, 1991; Malliaris, 1994) suggests that nonlinear models capture market processes better than linear models. The current study tests the ability of complex nonlinear modeling techniques to capture the relationship between accounting variables and subsequent stock returns.

The complex nonlinear modeling techniques used are neural networks and dual-layered genetic algorithms. The third section reviews the use of neural networks in market research. The studies reviewed normally compare the neural network results with linear models. These results indicate the promise of neural networks to map the complex markets better than traditional methods. Adaptive algorithms like neural networks can address the problem of market nonlinearity. The modeling technique that is most capable of appropriately capturing this market complexity should generate the largest excess returns.

According to Matthews (1996) neural networks are "notoriously opaque" in their reasoning. Genetic algorithms address this opaqueness and are included in the current study. The use of genetic algorithms and associated rule induction methodologies in market studies (Ruggiero, 1995; Ruggiero, 1997a; Fogler, 1995; Colin, 1994; Edmonds, et al., 1995; John, et al., 1995; Lam et al., 1995; Kiernan, 1994) are reviewed. The dual-layered genetic algorithm combines the use of rule induction methodologies with genetic algorithms to improve performance. The major advantage of this approach over neural networks is the ability to comprehend the trading rules. This research is more practically oriented and many of the studies are proprietary.

The final section discusses the possible effect environmental factors have on the relative importance of accounting information. This study looks at the inclusion of environmental factors through an industry proxy. Prior research suggests that information varies in importance across industries (Downs, 1991; Biddle and Seow, 1991; Lustgarten and Thomadakis, 1980; Ely, 1991; Clarke, 1989). If information varies by industry then industry-specific models may be more appropriate than a market-wide model.

2.1 Fundamental analysis

In equilibrium there must be enough mispricing to provide incentives for the investment of resources in security analysis (Palepu, et. al., 1996). Since financial analysts invest a substantial amount of time and resources in analyzing securities there must be some mispricing. Financial analysts use the accounting information reported in the financial statements to obtain insights into the economic activity of business entities (Palepu, et. al., 1996), as part of the security analysis.

This economic activity of an entity helps to create value by generating returns in excess of the cost of capital. Financial analysts try to determine which firms will generate greater value than normal. Bernard and Thomas (1990) provide evidence that the initial reaction of stock prices to news tends to be incomplete. This incomplete reaction provides support for the use of a trading strategy to outperform market averages. The remainder of this section reviews research suggesting the existence of mispricing and discusses the use of accounting information to locate this mispricing.

One possible reason that the market does not incorporate the full effect of news into the market price immediately is the inability to properly determine the risk associated with the firms' news. Chan and Lakonishok (1992) reviewed beta, the traditional assessment of a firm's risk, and found no conclusive evidence in support of beta as a measure of risk. They examined the entire CRSP history of stock returns and found no strong support for beta. One possible explanation as to why beta is not supported is because information is not distributed in a timely fashion to the potential investors. If this explanation is correct then the more recent time periods should provide stronger support for beta due to an increase in the information distribution efficiency resulting from technology improvements. Cahn and Lakonishok's (1992) results do not support this explanation. The most recent time periods show little to no support for beta. Fama and French (1992) found that beta does not seem to help explain the cross-section of average stock returns. Beta-sorted portfolios do not support the Sharpe-Lintner-Black (SLB) asset-pricing model while size-sorted portfolios do. Beta-sorted portfolios do not help explain average stock return of 1963-1990.

A number of other anomalies indicate that mispricing exists. Fama and French (1992) have suggested that the combination of size as measured by market equity (ME) and book-to-market equity (BE/ME) absorbs the roles of leverage and earnings-to-price (E/P) in average stock returns. This may occur because stock risks are multidimensional, with one dimension proxied by ME and another by BE/ME.

Chan, Hamao, and Lakonishok (1992) found a significant relationship between earnings yield, size, book-to-market ratio, cash flow yield, and expected returns in the Japanese market. They found that the book-to-market ratio and cash flow yield had the most significant positive impact on expected returns. A negative relationship was found to exist between equity returns and size.

French (1980) found that Monday returns are consistently negative. He claimed that it is hard to imagine an equilibrium consistent with both market efficiency and negative expected returns. Keim and Stambaugh (1984) found that the weekend effect is strong no matter what time period or firms. The effect occurs whether the weekend is one or two days. Abraham and Ikenberry (1994) suggest that the negative returns on Mondays might be the consequence of returns in prior trading sessions. They suggest that individual investors use the weekend to gather information and sell stocks Monday morning and that individual investors are more active sellers of stock on Mondays. This confirms the idea that not all information is compounded into the price immediately. Consequently, this argument could lead to continuing returns from fundamental analysis due to the individual investor's delay in processing the informational content of the financial statements. Agrawal and Tandon (1994) found that this weekend effect applies to markets around the world.

Keim (1983) shows evidence that daily abnormal return distributions in January have large means and the relation between abnormal returns and size is always negative and more pronounced in January. The magnitude of the size anomaly does not appear sensitive to estimators of beta. Tax loss selling and information hypotheses do not appear able to explain the anomaly. A risk proxy that only applies one month of the year does not seem consistent with the efficient markets hypothesis. The January effect is not confined to small firms (Agrawal and Tandon, 1994).

Abnormally high returns on the trading day before holidays exist in the U.S., the U.K., and Japan (Kim and Park, 1994). The holiday effects in the U.K. and Japan are independent of the holiday effect in the U.S. Agrawal and Tandon (1994) found this holiday effect in a number of other countries.

Kramer (1994) suggests that a multifactor model with seasonal risk may account for the January effect found in market studies. He argues that expected-returns shifts are responsible for the January seasonals in low-priced company returns. This argument is not consistent with the international findings. Gangopadhyay (1994) found that the risk-return relation to be applicable only in January. Because transaction costs are higher for small firms and, the January effect is more pronounced for small firms, Schwert (1989) included transaction costs in an attempt to explain the effect. He found that controlling for transaction costs does not explain all of the size effect in Australia and the US. He found that the magnitude of the size effect varies over time and is related to Price/Earnings ratios.

Agrawal and Tandon (1994) examined the day-of-the week effect, turn-of-the-month effect, holiday effect and month of the year effect for eighteen countries: Belgium, Denmark, France, Germany, Italy, Luxembourg, Netherlands, Sweden, Switzerland, UK, Hong Kong, Japan, Singapore, Brazil, Mexico, Canada, Australia, and New Zealand. The authors claim that "given the existence of these anomalies worldwide, one conclusion seems warranted: we cannot attribute them merely to noise, data snooping or selection biases."

Hand (1990) tested a hypothesis he called "extended functional fixation hypothesis." Under the hypothesis a marginal investor determines a firm's price. This marginal investor does not have to be sophisticated as required under the efficient market's hypothesis. Hand shows that the extended functional fixation is appropriate using debt-equity swaps. Because of the extended functional fixation hypothesis effect on the efficient markets' hypothesis, Ball and Kothari (1991) attempted to show that Hand's results were nothing more than a proxy for size effects that are a proxy for risk. Hand (1991) showed that his results were not a proxy for risk.

Wang (1993) looks at the anomaly of excess price volatility in the U.S. and the U.K. The U.S. and the U.K. stock prices are more volatile than can be justified based on expected future cash flows generated by stocks. A possible explanation is that a group of agents follow a rule of thumb decision-making procedure. This suggests that near-rational behavior may be an important force on financial markets. Steiner and Wittkemper (1994) look at the "Coherent Market Hypothesis," which drops the premise of rational investors and thereby relaxing the precondition of approximately normally distributed stock returns. The authors state "common approaches for explaining the relationship between asset prices, risk and return like the CAPM, APT and Sharpe's market model are based on unrealistic assumptions and allow only a vague and coarse modeling of the real dynamics of the capital markets." The introduction of near-rational agents suggests that rational investors could systematically out-perform the market.

Chan, Hamao and Lakonishok (1992) also found a significant relationship between book to market ratio and expected returns in the Japanese market. They also found significant relationships between size and cash flow yields and expected returns.

Ali (1994) shows that earnings, working capital and cash flows contain incremental information using a nonlinear model. Prior research using linear models has not provided conclusive evidence on the incremental information content of working capital and cash flows from operations. Finger (1994) found evidence that earnings are not a random walk for four and eight year periods.

Ball (1992) has suggested size as a proxy for risk. Ball claims that "the financial-statement-information anomaly appears due to accounting ratios proxying for stocks' expected return." He also argues that size is a good predictor of expected returns and that it does not challenge the efficient market theory. Chan, Hamao, and Lakonishok (1992) looked at the effect of size in Japanese markets and found that it was dependent on the specific model and time period. Thus it is not as robust in the Japanese markets. Shevlin and Shores (1993) suggest that firm size may be proxying for some misspecification of the relation between cumulative abnormal returns and unexpected earnings. The coefficient for firm size is negative for observations with positive unexpected earnings and is positive for observations with negative unexpected earnings.

Earnings changes and levels have been related to excess returns by a number of researchers (Easton, 1985; Collins and Kothari 1989; Ahmed 1994; Ali 1994; Finger 1994; Bhushan 1994; Ball 1992; Ali and Zarowin 1992; Easton and Harris 1991). Easton (1985) provides empirical evidence of an information link between accounting earnings and future dividends and a valuation link between future dividends and security prices. This relates earnings to returns through the present value of expected future dividends. Collins and Kothari (1989) empirically show that the earnings response coefficient (ERC) has a positive relation with market-to-book value of equity. Ahmed (1994) found evidence of unexpected earnings reflecting financial information about future economic rents.

Previous attempts to develop a trading strategy based on specific accounting information or 'fundamental analysis' have generated excess returns. Ou and Penman (1989) empirically examined 'fundamental analysis' and used it to generate abnormal returns. They used sixty-eight accounting attributes and stepwise logistic regression to predict positive or negative changes in earnings. Using logistic regression they computed the probability of positive earnings' changes for individual firms, P_r. They then employed a trading strategy that required buying long, firms with high P_r values and selling short, firms with low P_r values. Their model earned excess returns of 8.3 percent for the 1973-1983 period.

Holthausen and Larcker (1992) conducted an analysis similar to that performed by Ou and Penman but used subsequent excess returns as the binary dependent variable instead of excess earnings. Stepwise logistic regression was used to predict positive or negative changes in excess returns. They then used a trading strategy similar to Ou and Penman's (1989), buying long, firms with a high probability of having positive excess returns for 12 months and selling short, firms with the lowest probability of having positive excess returns. This generated a more profitable trading strategy than using the earnings changes as the dependent variable. For the 1978-1988 period their strategy earned a 4.3 percent excess return. Using Ou and Penman's P_r measure the same excess return measure was -.01 percent.

Holthausen and Larcker modeled excess returns directly. They reason that this should allow for better prediction of excess returns than when a proxy for excess returns is used. Ou and Penman used the change in earnings as a proxy for the change in excess returns. The noise of the proxy measure reduces the predictive ability of the model. Using the change in earnings establishes a more direct link to the accounting information. Holthausen and Larcker were primarily concerned with the possibility of earning abnormal profits rather than further establishing the link between accounting information and market values.

Ou and Penman tried to establish a link between accounting information and market returns. They used all feasible accounting attributes to find the most useful measures. This resulted in sixty-eight attributes being used initially. Holthausen and Larcker started with the same attribute set as Ou and Penman. More recent studies stress the quality of information over the quantity. Lev and Thiagarajan selected attributes based on theoretical justification. This study incorporates the attributes with theoretical justification and those attributes which Ou and Penman and Holthausen and Larcker found significant. This reduces the initial attribute set, which is important in the development of complex nonlinear models.

The theoretical effect of risk on returns is well established. If the future stream of returns is uncertain then the value will be lower than if the stream of returns is certain. One problem with incorporating risk into modeling is that it requires looking into the unknown future. The amount of uncertainty a firm’s earning stream possesses is not a measurable concept at this point. The measures of Holthausen and Larcker (1992) and Ou and Penman (1989) have been suggested as a good risk proxy. Chan and Lakonishok offer an alternate explanation to these items proxying for risk (1992): " . . . behavioral and institutional factors, unrelated to risk, play a major role in generating stock returns...." These items include environmental factors.

In summary it appears that excess returns can be found by using fundamental values. This attempt to link fundamental values to excess returns is a form of `fundamental analysis'. Since risk is an ill-defined concept, any relationships between fundamental values and excess returns may be capturing risk. This issue cannot be controlled for until risk is properly defined.

2.2 Problems of Complexity and Nonlinearity in Market Research

Research has shown that linear models may not always be the most appropriate forms for modeling the relationship between financial statement data and market data. Freeman and Tse (1992), show that a nonlinear, arctan model can explain the differences between the earnings response coefficient and the price-earnings coefficient better than the normal linear models. A linear model assumes that the response of firms to unexpected earnings is constant among the ranges of unexpected earnings. They state that "the absolute value of unexpected earnings is negatively correlated with earnings persistence." This suggests that one can classify earnings' surprises between permanent and transient, and that permanent unexpected earnings have a greater impact than transient changes. Their regression results suggest that nonlinear models have greater explanatory power.

Zapranis and Refenes (1994) state that "there are strong reasons to believe that the relationships between asset prices and their determinants are determined by a complex nonlinear process. Neural networks provide a suitable methodology for modeling this type of relationship." Zapranis and Refenes tested their assumptions about the suitability of neural networks to model complex nonlinear market processes using tactical asset allocation. They compared the performance of linear regression against a feedforward ordinary least squares network. The authors used stepwise regression to select the variables. They then developed neural networks for the same time periods using the variables selected by the stepwise regression procedure. The performance measure for model comparison was R^2. They used five different models and in all cases the R^2 for the neural network was significantly better than the R^2 for the stepwise regression models.

Brorsen and Yang (1994) found that the empirical distribution of daily price changes has more observations around the mean and in the extreme tails than a normal distribution. Scheinkman and LeBaron (1989) found evidence of nonlinear dependence in stock returns. They used weekly return series from the value-weighted portfolios of the Center for Research in Security Prices at the University of Chicago (CRSP). They tested the variation on the weekly returns to determine if it was more feasible to generate it from randomness or nonlinearities. They concluded: “a substantial part of the variation on weekly returns is coming from nonlinearities as opposed to randomness.” Violating the independence assumption causes ordinary least squares estimates of regression parameters to be inefficient. Brock, Hsieh, and LeBaron (1991) suggest that asset returns may not follow a stochastic process but might be generated by deterministic chaos.

These findings of complexity and nonlinearity have been primarily limited to US and London markets. Corhay and Rad (1994) reviewed stock returns in other markets to see if these nonlinearities held. They examined stock indices in France (CAC General), Germany (Commerzbank), Italy (Milan Banca), the Netherlands (ANP-CBS general) and UK (FT All-Shares) to see if these nonlinearities held in smaller markets. Corhay and Rad (1994) found that all distributions are negatively skewed and exhibit high levels of kurtosis. Using the Kolmogorov-Smirnov D-statistic, they reject the normality assumption for all cases, along with homoskedasticity. These results indicate complexity of foreign markets as well. Corhay and Rad (1994) state that the European indices “exhibit nonlinear dependence." Specific testing for nonlinearity on foreign markets remains to be completed. This suggests that undefined, complex nonlinear phenomena are universal.

Malliaris (1994) suggests there is a non-random underlying structure to market returns labeled chaotic dynamics. Chaotic dynamics is a deterministic model that yields a time series behavior that appears random when in fact a nonlinear deterministic equation generates it. According to Malliaris, support for the chaotic paradigm implies that active management of an S&P portfolio is possible. Scheinkman and LeBaron (1989) imply that the S&P follows a nonlinear deterministic model, which is consistent with chaotic dynamics. This supports the idea of earning excess returns through fundamental analysis, if the model developed captures the nonlinear deterministic equation. Hsieh (1991) says that chaotic dynamics can generate large movements that appear random, with greater frequency than linear models. Chaotic behavior can potentially explain fluctuations in the economy and financial markets that appear random. If a not-too-complex chaotic process governs markets, it should have short-term predictability. Traditional linear forecasting methods would not work; nonlinear models must be developed. Hsieh strongly rejects the hypothesis that stock returns have independent and identical distributions (IID) for the market as a whole and for decile portfolios. The weekly S&P returns and the daily returns in 1982-1989 and the fifteen-minute returns in 1988 are also not IID. Hsieh claims that the rejection from IID is not due to low complexity chaotic behavior in stock returns but is due to conditional heteroskedasticity. These studies have mixed results about the presence of chaotic behavior. This suggests that some complex deterministic process may allow excess returns to be generated by fundamental analysis.

2.3 Neural Networks in Market Studies

Neural networks are frequently compared to traditional statistical models in current market studies. The neural network models outperform the traditional models in many of these studies. This section reviews some studies that compare the performance of neural networks with the traditional statistical models. In addition to academic studies many investment agents are using these advanced quantitative tools. Shirreff (1994) indicates that Fidelity Management and Research, The Burney Co., and Kleinwort Benson Investment Management were using neural networks and genetic algorithms in trading by 1994. Fogler (1995) suggests neural network architectures are totally generalizable. Neural networks can also encompass formulations that include clustering and fuzzy sets as well as nonlinear problems like those found in chaos.

Marc Levitt (1994) used backpropagation neural network and genetic algorithm based learning techniques to generate abnormal returns in options markets. The author applied a trading strategy developed from the techniques to buy and sell options for 1993. The best cumulative return using these advanced techniques for 1993 was 43 percent. The daily and monthly returns have a 99.5 percent probability that they are greater than zero. This result indicates that these advanced techniques may generate excess returns in the stock market.

Zapranis and Refenes (1994) compare the performance of neural networks with stepwise linear regression in forecasting the differences in returns between equities and cash versus estimated differences in returns between bonds and cash. They used monthly economic variables to estimate differential returns for one month ahead. They biased the results towards the linear regression by only using the variables deemed significant in the stepwise regression as input nodes in the neural networks. The neural network models outperform the multiple linear regression models in terms of R^2.

Haefke and Helmenstein (1994) compare the performance of linear regression models and neural networks using the Austrian Traded Index (ATX). The ATX forecast using neural networks produces better results than the linear regression model. Using a Theil[2] measure, the linear regression model performs worse than the random walk model, while the neural network model performs slightly better than the random walk model.

Bolland and Refenes (1994) compare the performance of linear regression with neural networks in modeling the relationship between volume and open interest with futures prices. The networks out-performed linear regression in modeling the relationship between volume, open

[2] Thiels measure is based on the RMS scaled to fall between 0 and 1. For a more complete discussion on this measure see Pindyck, R. and D. Rubinfeld, 1981, <u>Econometric Models & Economic Forecasts</u>, McGraw Hill, New York, pp. 360-67.

The calculation is on the next page:

$$\frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^a)^2}}$$

Y_t^s = *forcasted value of* Y_t
Y_t^a = *actual value of* Y_t
T = *number of periods*

interest, and futures prices as measured by RMS[3] and R^2.

Hall (1994) uses neural networks to manage a stock portfolio for the Pension and Investment Department of Deere & Company. He modeled the predicted change in a stock's price relative to the Standard and Poor's 500 average. Hall found that between four and eight inputs are required to estimate the future price of an individual stock. The portfolio's performance has exceeded the S&P 500 since inception after all transaction costs.

Yoda (1994) developed a neural net-based prediction system for the Tokyo stock market. The neural network predicts the timing of buy-and-sell signal for stocks on the Tokyo stock exchange price index. The system was used for the test period from September 22, 1989 to October 22, 1992. The per annum return for this time period was 1.60 percent compared to a buy-and-hold strategy per annum return of -20.63 percent.

Ye and Gu (1994) developed three different neural networks for trading on the Shanghai Stock market. The networks were for an ascending trend, a descending trend and a stationary environment. Using the appropriate model helped them to have a correct prediction rate for price changes of 92 percent in the training set and 74 percent on the test set.

Gencay (1995) compared the performance of ordinary least square regression (OLS) with a single layer neural network regression model to determine the appropriate methodology to model stock returns. The models forecasted the change in the Daily Dow Jones industrial Average and the results were compared to a moving average benchmark. The network model had an average of 12.9 percent reduction in the Mean Square Error over the OLS regression. Gencay states that this additional explanatory power supports the nonlinear predictability in stock market returns.

Robles and Naylor (1995) examined the performance of neural networks in COMEX trading from 1989 to 1993. They developed artificial neural networks (ANN), a weighted moving

[3] RMS is a measure of the forecast error. For a more complete discussion on this measure see Pindyck, R. and D. Rubinfeld, 1981, <u>Econometric Models & Economic Forecasts</u>, McGraw Hill, New York, pp. 360-67. The calculation is on the following page:

average model, and a Martingale process to compare results with a buy and hold strategy. They developed the models using the time period from December 1, 1989, to April 12, 1992, and then used the period from December 7, 1992 to November 29, 1993 as the test period. The results indicate that the ANN models the process best. The average return using the ANN was 63 percent greater than with the weighted moving average and 36 percent greater than the Martingale process.

Bentz, Refenes and De Laulanie (1995) compared the performance of a linear regression with a neural network. The authors state that neural networks' "powerful universal approximation abilities allow them to model nonlinear functions, in particular conditional relationships." The authors used financial data on ninety French stocks for the blue chip French stock index, SBF250, along with some economic indicators to develop investment strategies based on linear regression and neural networks. The data were monthly from January 1988 to January 1995. The neural network model had a better fit in the training period prompting the authors to claim the "difference in fit between linear and nonlinear models suggest that some nonlinearities exist in the data set." The profits generated from the neural network based model were consistently higher than the profits from the linear regression model. The authors included transaction costs of one percent.

Yao and Poh (1995) used neural networks to model the Kuala Lumpur Stock Exchange (KLSE). They attempted to buy and sell the KLSE index, which consists of eighty-six major Malaysian stocks, on a daily basis. The annual return using six different network structures varied from nine percent to 26 percent using the hold out sample. This result supports the argument that neural networks can be used to generate positive returns in a trading strategy.

Avouyi-Dovi and Caulet (1995) compared the performance of neural networks to traditional statistical methodology in predicting the direction of daily change in financial

---

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \qquad \begin{aligned} Y_t^s &= \text{forcasted value of } Y_t \\ Y_t^a &= \text{actual value of } Y_t \\ T &= \text{number of periods} \end{aligned}$$

instruments. They compared the performance of a neural network with an ARMA[4] (1,1) and an AR[5] (3) model for the German Mark-Dollar exchange rate, the French Franc-German Mark exchange rate, CAC index, S&P 500 index and DAX index. The daily data was from October 3, 1988, to June 12, 1995. During this time frame the neural network predicted the correct change better than the ARMA (1,1) and AR (3) for every instrument except the French Franc-German Mark exchange rate. The prediction accuracy rate for the neural network was as high as 92 percent for the German Mark-Dollar exchange rate while the ARMA (1,1) model never had over a 50 percent accuracy rate and the AR (3) model's best performance was a 53 percent accuracy rate. The authors found that by trading based on these predictions the neural network averaged a return of 24 percent for all five instruments while neither of the other methods generated an average return over one percent.

Chiang, Urban and Baldridge (1996) compare the performance of neural networks with traditional econometric techniques including linear and nonlinear regression. They forecast the end-of-year net asset value of mutual funds and show that the neural networks significantly outperform regression models with limited data.

Donaldson and Kamstra (1996) used time series forecasts for market data from Canada, Japan, the UK, and the US. Their results indicate that neural networks have the best performance for these international markets. They also point out that over fitting causes the performance of neural networks to deteriorate.

Ruggiero (1994) claims that neural networks are good for the following trading-related applications: predicting the future direction of a commodity; selecting between trading systems; predicting the trend of a commodity; making subjective analysis mechanical and predicting tops

[3] Autoregressive-Moving Average Model of Order p and q (ARMA (p,q)) is a combination of an Autoregressive Model of order p and Moving Average Model of order q. The equation is as follows:

$$y_t = f_1 y_{t-1} + f_2 y_{t-2} + \ldots + f_p y_{t-p} + q_0 + e_t - q_1 e_{t-1} - q_2 e_{t-2} - \ldots - q_q e_{t-q}$$

[5] Autoregressive Model of order p AR(p). For a complete discussion on the ARMA(p,q) and AR(p) models see N. Farnum and L. Stanton, 1989, Quantitative Forecasting Methods, PWS-KENT, Boston, MA, pp. 446-451. AR(p) is calculated as follows: $y_t = f_0 + f_1 y_{t-1} + f_2 y_{t-2} + \ldots + f_p y_{t-p} + e_t$

and bottoms. He developed a multiple neural network model to predict the Standard and Poor's 500 for a five-week period. The system generated returns of 15 to 30 percent.

2.4 Genetic Algorithms in Market Studies

Genetic Algorithms are in the developmental stage in market studies. Rule tree induction and neural networks have been used to develop models from databases. These methods have met with limited success and comprehension. This has lead to a recent adoption of genetic algorithm techniques.

Ruggiero (1995) argues that “genetic algorithms are a simple but powerful tool for finding the best combination of elements to make a good trading system or indicator and are especially good in evolving rules to create artificial traders.” Ruggiero (1997a) indicates that developing a good trading strategy is an iterative process that may require weeks or months for the best market analysts. The genetic algorithm approach can speed up the development by “a factor of 1,000 to 10,000 or more.” Fogler (1995) indicates that because investment problems have decision variables that are both discrete and correlated, the optimal answer is combinatorial. Genetic algorithms are suited to handle complex combinatorial problems.

Colin (1994) was a manager with the treasury department of Citibank in London where genetic algorithms were used for currency trading. Colin determined that genetic algorithms were useful tools in trading and Citibank established a fund that successfully traded on the rules generated by the genetic algorithms.

Edmonds, Brukhardt and Adjei (1995) used genetic programming and fuzzy operators to generate excess profits in financial trading. The authors developed production rules from genetic programming for the trading of financial instruments. These instruments included US Treasury Bonds, NIKKEI index, FTSE index, Standard and Poor's index and the Deutsche Mark to Dollar exchange trading. Each instrument was examined for a five hundred-day period. Four hundred and seventy days were used for training and thirty days for testing. In all cases the production rules generated a profit in excess of the related instruments' return. The Deutsche Mark to Dollar exchange trading generated the greatest out of sample return a 13.73 percent annualized profit. Profits were accumulated and not committed back into the instruments. They also used a stopping condition of five percent down from peak price.

John, Miller and Kerber (1995) used a rule induction methodology to generate trading rules using knowledge discovery techniques. These techniques generated a trading rule set on a quarterly basis. The data set used included over one thousand companies and the portfolio chosen consisted of the highest ranked fifty stocks each quarter. Using a one percent trading cost over a four-year period the rule set had a total return of 238 percent compared to market-capital-weighted portfolio of all stocks in the universe of 93.5 percent. The authors also used an out-of-sample test period in which the rule induction method generated a return of 32 percent to a benchmark return of 12 percent.

Lam, Chiu and Chan (1995) used a series of if-or-then rules to develop a trading strategy on the Heng Seng Index. They used six blue label stocks and determined rule sets to predict to buy or sell the individual stocks. The use of if-or-then rules is a fuzzy knowledge based approach. The authors found an 88 percent accuracy rate for predicting the correct buy/sell signal for the time period from January 1994 to April 1995.

Ruggiero (1997b) developed trading strategies in two markets using genetic algorithms. The first market was T-Bond trading. The development time period was from January 2, 1987, to December 31, 1994, and the testing period was from January 1, 1995, to November 25, 1996. Implementing two rules developed in the training period for T-bonds during the testing period resulted in sixteen trades of which 75 percent were profitable. The second market study was the Swiss Franc-Dollar exchange. The development period was from February 13, 1975 to January 31, 1994, and the testing period was from January 1, 1995, to November 25, 1996. The results of two rules show that profits were earned in 64 percent of the trades.

Matthews (1996) reported that the Thomas J. Watson Research Center developed a genetic algorithm model for trading stocks on the American markets. The Research Center used 40 financial indicators, such as average monthly earnings and investment opportunities, to develop a trading rule set. They found trading rules, using a proprietary methodology, that generated a 270 percent return for a five-year period when the related market return was 110 percent.

Kiernan (1994) surveyed businesses to determine if advanced algorithms were being used in practice. He noted that First Quadrant uses genetic algorithms to help manage five billion dollars of investments. The company reported that the genetic algorithms were directly responsible for twenty-five million dollars in revenue in less than a two-year period. David Leinwebber the Director of Research for First Quadrant states that genetic algorithms “allows us to refine ideas we know are sensible in a general sense.” The results from the US markets were so promising that the firm used the genetic algorithms in seventeen different countries to develop investment strategies. Richard Bauer, another consultant interviewed, stated that genetic algorithms are “the tools, the mental swords, used by those who want to do business in today’s financial markets.”

2.5 Environmental factors

Industrial organization economists have employed the "structure-conduct-performance" paradigm since its introduction by Bain (1956). This paradigm alleges that firm performance depends on the conduct of buyers and sellers, which depends on market structure characteristics such as barriers to entry, product type, and the degree of vertical integration. Porter (1980) describes “five forces”: rivalry among existing firms, threat of new entrants, threat of substitute products, bargaining power of buyers, and bargaining power of suppliers that affect industry profitability. Downs (1991) found that the ratio of market value to the fair value of assets differs widely across industries.

Biddle and Seow (1991) show that the earnings response coefficients vary across industries. For example manufacturing industries exhibit larger response coefficients than utilities and food and kindred products. They removed all firms with missing data from Standard and Poor’s Compustat industrial files and CRSP monthly returns master file. The remaining 659 firms were then grouped into forty industries. The industry groups were arranged by looking at the four digit codes and similar size groups. There were six two-digit codes that were not grouped in the same industry classification. In each of these cases there were a large number of firms within the two-digit classification and were therefore reduced into smaller groupings. Biddle and Seow (1991) show that earnings response coefficients are negatively related to financial and operating leverage and positively related to growth, product durability, and industry barriers to entry. Their results indicate that environments are related to industry through the Standard Industrial

Classification (SIC) code.

Specific industries should have unique financial information critical for predicting future performance. Greig (1992) claims that Ou and Penman (1989) introduce industry specific results by using specific ratios. The values of the ratios differ by industry, thereby factoring in industry-specific information. However, different ratio values across industries may cause ratios to be excluded from a market-wide model that may be critical for a specific industry. Lustgarten and Thomadakis (1980) found that stock market reactions to earnings changes are greater for firms in industries with fewer firms. Ely (1991) assumes those firms in the same industry face similar production environments, including technology, product markets, and regulatory environments. Ely defines industries by the SIC codes. Four industries are used. Ely shows that there are inter-industry differences in the types of management compensation plans and the relationship between compensation and accounting based firm performance measures.

Clarke (1989) hypothesized that firms classified by SIC should display similar sales changes, profit rates, and stock price changes. Clarke tested this hypothesis using the SIC industry code, proxied by COMPUSTAT's DNUM, seasonal explanation, and yearly time trend variables as independent variables and sales changes, profit rates and stock price changes as dependent variables. Clarke ran regressions at the individual firm level, aggregated at the four-digit SIC code, the three-digit SIC code, the two-digit and one-digit SIC code. The most explanatory power was at the individual firm level; however the nonlinear modeling techniques require more observations than possible at the individual firm level. Biddle found no significant loss of explanatory power from the four-digit SIC code to the three-digit SIC code and from the three-digit SIC code to the two-digit SIC code. However, the study shows that there is a significant loss of explanatory power from the two-digit SIC code to the one-digit SIC code. Since the nonlinear modeling techniques require more observations than logistic regression some combination of firms is necessary. Because no additional explanatory power is lost until going from the two-digit SIC code to the one-digit SIC code the current study uses the two-digit code proxied by COMPUSTAT's DNUM for industry classification.

The first section reviewed the literature that indicates mispricing may exist. If mispricing exists then an appropriate trading strategy could generate excess returns. One of the most important steps in this study is to select the information that is not being fully utilized in market pricing. The literature in the first section indicates that a guided search, indicated by Lev and Thiagarajan (1993), of the accounting information could capture this mispricing. The next concern is to develop an appropriate model from this information. The literature in the second section suggests that complex nonlinear models may be required to capture this mispricing. The third section reviews the evidence supporting neural networks as an appropriate modeling tool in financial markets and supports its inclusion in this study. The fourth section discusses the benefits of using genetic algorithm based models to explain the market phenomenon and supports its inclusion in the current study. The final section reviews literature that supports the assumption that the type of information needed to capture this mispricing varies by industry. Market-wide models should be compared to industry-specific models to determine if this is correct.

CHAPTER 3
MODEL DEVELOPMENT

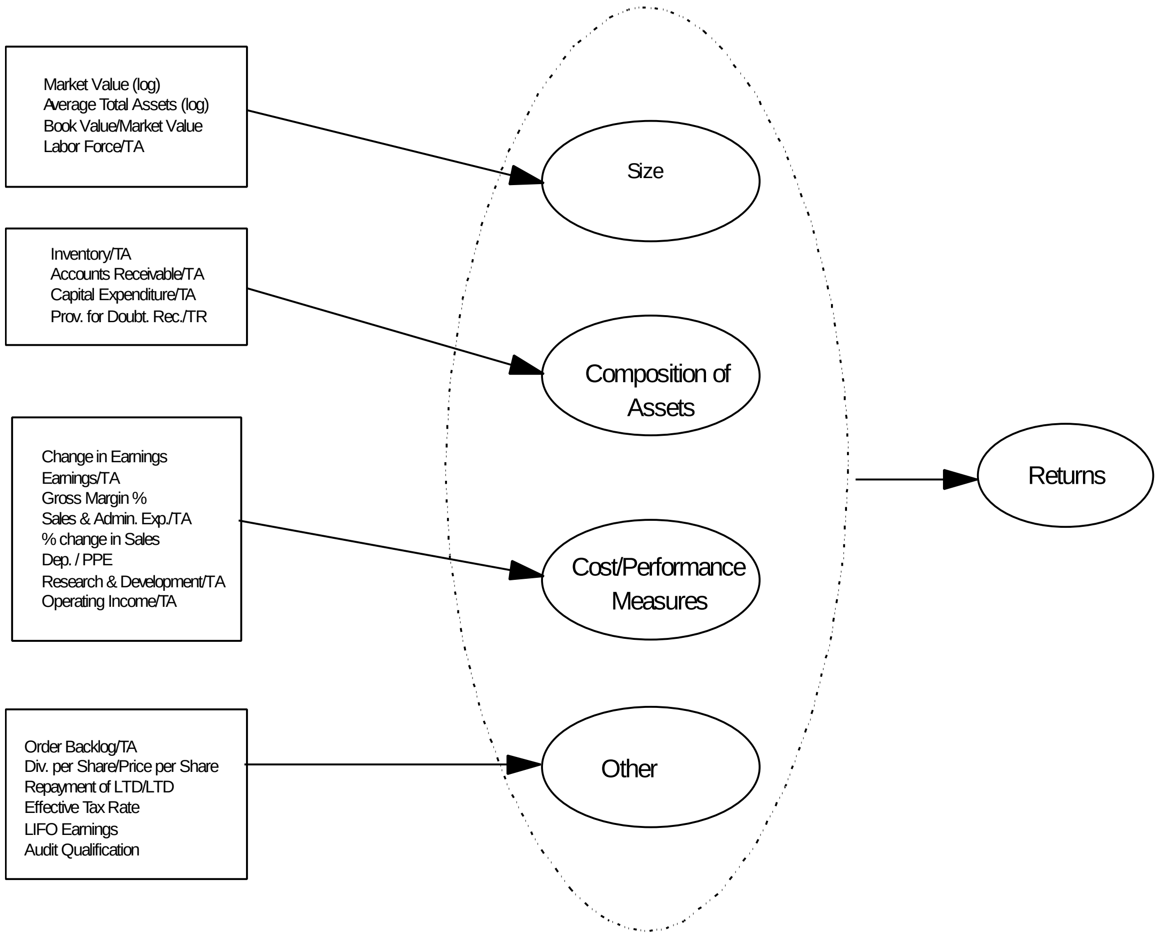
3.0 Introduction

Ou and Penman (1989) and Holthausen and Larcker (1992) have shown that accounting information can be used to systematically generate abnormal returns. This study investigated the use of complex nonlinear modeling techniques to generate abnormal returns. As in Ou and Penman (1989) and Holthausen and Larcker (1992), transaction costs were not considered. This affects the true profitability of the trading strategy, but allows for comparisons with Ou and Penman and Holthausen and Larcker's results. The basic model is shown in figure 1. (Next page)

Section 3.1 describes the independent variables included in the models' candidate list. Independent variables are selected using a guided search. The candidate variable selection is primarily based on the empirical findings of Ou and Penman (O&P), Holthausen and Larcker (H&L) and the analysts' discussion in Lev and Thiagarajan (L&T). Section 3.2 describes the dependent variables. The dependent variable is a 0,1 binary variable. The dependent variable is set to one if abnormal returns for a prior twelve-month period are positive and zero otherwise. Two different return measures are used; this is consistent with Holthausen and Larcker (1992). The logistic regression models are discussed in section 3.3. These models are used as base models for comparison with the complex nonlinear techniques.

Hypothesis development is reviewed in section 3.4. The first proposition evaluated was that complex nonlinear modeling techniques represent the relationship between financial data and market returns better than logistic regression. The complex modeling techniques examined were neural networks and a dual-layered genetic algorithm approach that used classification trees and genetic algorithms. The returns generated from a trading strategy based on these techniques were compared to returns based on logistic regression. The final proposition studied is that the financial variables have varying impacts depending on a firm's present environment or industry. The final section discusses the data collection.

Figure 1: Basic Model



3.1 Independent Variables

The independent variables were selected from prior empirical evidence and logical arguments. Variable selection was primarily based on three studies. The first group of variables considered in this study were the six variables found significant in both of the time periods examined by Ou and Penman. These variables are: percentage change in inventory/total assets, change in dividends per share, percentage change in capital expenditures/total assets lagged one year, return on total assets, operating income/total assets, and repayment of long term debt as a percentage of total long term debt. The next set of variables considered are those variables that entered four or more of the twelve Holthausen and Larcker models. These are: percentage change in total assets, depreciation/property plant and equipment, percentage change in depreciation/property plant and equipment, gross margin, dividends per share and percentage change in sales. These variables have empirical support for their inclusion in the model building phase.

The next set of variables examined in this study is based on Lev and Thiagarajan (1993). Lev and Thiagarajan defined and tested twelve fundamental signals, seven of which were significant at the .05 alpha level in their restricted sample. These signals are: inventory, accounts receivable, capital expenditure, gross margin, sales and administrative expenses, order backlog and labor force. Since Lev and Thiagarajan logically justified all of their signals, measures for signals without empirical support were candidate variables in the current study. These signals are: research and development, provision for doubtful receivables, effective tax rate, LIFO earnings and audit qualifications. Fama and French (1992) have suggested market value and the ratio of book value to market value as risk proxies and these are also included in the variable list. The final variable considered was earnings changes, which several researchers have related to returns (Easton, 1985; Collins and Kothari, 1989; Ahmed, 1994; Ali, 1994; Finger, 1994; Bhushan, 1994; Ball, 1992; Ali and Zarowin, 1992; Easton and Harris, 1991).

The prior studies provided support for twenty-two variables. These variables were classified into four major groupings. The first major grouping addressed size problems inherent in market studies. Bernard and Thomas (1989) stress the importance of size in market studies. The next major grouping was based on the ability of firms to respond to market conditions. The composition of firms' assets indicates how quickly firms can respond to changing conditions. This response time may impact investors' decision processes. The next group of variables addressed the economic performance of a firm. The historical economic performance of a firm can give insights into the effectiveness of a firm's management. The effectiveness of a firm's management should impact the market response to the firm. Cottle et. al. (1988) state that "the most convincing proof of capable management" is superior historical performance. The last group of variables includes signals that affect the firm's valuation but are not traditional accounting information, and include economic information.

3.12 Size Variables

The first size related variable is market value (MV). Market value is a measure of size that is consistent across industries. Market value is calculated by taking the log of the number of shares outstanding multiplied by the closing price per share at the company's year-end. This measure is independent of industry capital requirements.

The second size related variable is average total assets (ATA). Total assets are the log of the book value of a firm's assets. This measure of size while correlated with market value will yield a different ranking. Service firms, for example, require smaller capital investment than manufacturing firms, service firms with large market values may be grouped with small manufacturing firms that have similar total assets values. Total assets only reflect items that have a historical cost. Service industries have a large investment in personnel that does not show up in total assets since employee expenses are not capitalized.

Book to market ratio (BVMV) is another variable used. The book to market ratio considers the total amount of capital invested in a firm plus some industry information factored in through the amount of book value that a firm has recorded. This measure is an attempt to show the differences from the market value and the total assets. BVMV is a common measure that is aimed at trying to control for the inconsistencies in the total assets variable. The last size-related variable is labor force size divided by total assets (LF). Labor force size may be more appropriate for service industries that have few capital assets.

3.13 Composition of Assets Variables

The second set of variables is concerned with the composition of the firm's assets. The first variable in this group is inventory as a percentage of total assets (INV). Inventory is a critical component for many firms, particularly in the retail and wholesale industries. Managing the cost associated with inventory is key to firm profitability in these industries. In addition to the direct cost of inventory, firms have carrying costs and replenishment costs associated with the inventory. Excessive inventory levels will cause the carrying costs to be greater than necessary and reduce earnings available for returns to stockholders. A firm that tries to minimize its carrying costs through minimal inventory levels may incur losses of sales though shortages, which again reduce the earnings available for returns to stockholders. Low inventory levels may also increase the replenishment costs to an unacceptable level (rush orders for example). Proper inventory management will include minimizing these costs. Due to this impact of inventory management on earnings, some measure of inventory should be incorporated into the model. Inventory is dependent on size; therefore the measure of inventory should show the percentage of inventory to total assets.

Accounts receivable as a percentage of total assets (AR) is the next variable included. An accounts receivable percentage greater than the industry average could be an indicator of poor credit management. Conversely a percentage less than the industry average would suggest that the firm has an excessively stringent credit policy and is likely to lose profitable sales. Controlling for size differences, accounts receivables are scaled by total assets.

Capital expenditures as a percentage of total assets (FCE) is another composition of assets variable. Firms need to maintain technology and equipment at least to industry standards. This variable would detect firms that may be spending too much in a capital improvement plan or letting their capital deteriorate below competitive levels.

The final asset in this group is the provision for doubtful receivables (PDR). The percentage of receivables that are deemed uncollectible can also indicate how well credit is being managed. A percentage of doubtful receivables larger than the industry average indicates poor collection policies or poor authorization controls. A low percentage could again indicate that the credit approval is excessively strict and the firm is losing revenue. Both possibilities are indicators of poor credit management. Proper credit management is a good indicator of management's competence.

3.14 Cost/Performance Measures Variables

The next set of variables is cost/performance measures. The most obvious performance measure is a firm's earnings. Two aspects of earnings are significant in measuring a firm's performance, the direction of earnings as measured by the change in earnings (CE) and the earnings level as measured by earnings/total assets (ETA). The direction of earnings gives an indication of the probability of future earnings. If a firm has shown consistent earnings increases, then an investor would be more confident of that trend continuing than if the direction of earnings has been fluctuating. The level of earnings should be considered in conjunction with the direction of earnings. If the level of earnings is great then a decrease in earnings may not be a bad signal. It may be a signal that the firm was earning abnormal profits and competition forced a decrease in earnings to normal profits. An increase in earnings from a large deficit may not be a positive signal. The firm may not have reached a stable level of earnings, and may be have financial problems. Therefore measures for both the direction and level of earnings are included.

Components of earnings may also have good predictive power. The first component included is the gross margin percent (GMP). The gross margin provides a signal of the efficiency of a firm’s direct operations. A high gross margin suggests that the firm has good direct cost controls. This would be a positive signal for investors. The level of the gross margin needs to be based on the size of the firm for a meaningful comparison, thus gross margin percentage is used.

Sales and administrative expenses as a percentage of total assets (SAE) is the next component included. A large ratio could indicate that the firm has more layers of management than are required for maximum efficiency. A small ratio could indicate that the firm does not have adequate support staff to operate successfully.

The percentage change in sales (PCS) provides a measure of growth. If sales are increasing rapidly it is a signal to investors of the strength of the firm. This may be due to innovative advertising and marketing, or it may be from the production of a product superior to that of its competitors. Decreasing sales signal a weakness in the firm’s products or market and investors should be cautious of the firm.

Depreciation as a percentage of property, plant, and equipment (DPPE) is another signal of interest. This ratio indicates the age of the fixed assets of the firm in comparison with competitors. If the assets are fully depreciated then the ratio will be low, possibly indicating that there will need to be an infusion of capital to update the firm's plant capacity to remain a viable firm. If the ratio is high then the firm may be over investing.

Research and development costs (RD) are an indication of the firm’s commitment to remain competitive in an industry. Because the size of the firm affects the resources available for research and development, total assets deflate this measure. A large ratio may indicate that the company is improving its products in the market if the industry is growing. This improvement would increase the firm’s industry share and be a positive signal for investors. A low ratio may indicate that the firm is just going to continue with its current products but then lose market share to competitors who are investing in new product research. This would be a negative signal to investors.

Operating income as a percentage of total assets (OI) is a direct measure of the efficiency of firms' operations. Non-operating events affect the profitability of a firm but are not an indication of future operations. This allows investors to compare the future returns of various firms in their decision making processes. Operating income is dependent on size, as are all the financial variables, and must be divided by total assets to adjust for size.

3.15 Other Variables

The last set of variables is related to other economic events not recorded in the accounting information. The first item to include is order backlog (OB). This variable indicates prospective revenue for the upcoming periods. A large backlog may indicate that the demand for a firm's product will remain high. However, a large order backlog could be a negative signal if the firm operates in a highly competitive industry. In either case this variable signals future revenue from which investors will receive a return on their investment.

The next item is the dividend per share/price per share ratio (DSPS) that shows the return that the investors are currently receiving. Mature industries have to pay a higher rate of return to attract funds due to the lack of future growth opportunities. This also could indicate an over or undervalued firm to investors.

The next item included is the repayment of long-term-debt as a percentage of long-term-debt (RLTD). This item shows how much debt the firm is paying off currently: the greater the percentage of debt paid off the less cash required for interest payments in future periods. Repayment of long-term-debt is an important signal to investors because it increases cash available for return to the investors or for capital improvements.

A change in the effective tax rate (CETR) may indicate a problem with earnings persistence (Lev and Thiagarajan, 1993) that the market may incorporate into the firm's stock price. If investors feel that the earnings are not going to remain high then that will cause less interest in the firm. An unusual decrease in the effective tax rate is a negative signal to investors.

Management may show an increase in earnings by using an inventory cost method referred to as last-in-last-out (LIFO). When input prices are increasing the LIFO inventory valuation method assigns costs to the goods sold closer to the replacement value than other acceptable methods. The use of the LIFO cost flow assumption thus yields a conservative estimate of earnings and should be a positive signal to investors. The LIFO variable used in the study is a dummy variable, with a value of 1 if LIFO is used and 0 otherwise.

The audit qualification (AQ) indicates that the financial statements are not "fairly stated in all material respects." An unqualified audit report is the standard report issued by accounting firms. A qualification typically occurs when there is a problem with an accounting estimate or principle. In either case this indicates some doubt with respect to the firm's intentions. This is a negative signal to investors and is considered in this study as a dummy variable.

Table 1- Variable list

Variable	Formula	Reference
Size		
1) Market value(log) (MV)	(Close Price (24)*Common Shs.Out.(25))(log)	4
2) Average total assets (log) (ATA)	((Total Assets t-1 (6)+ Total Assets)/2)(log)	2
3) Book value/Market value (BVMV)	Book Value(37)/Market Value	4
4) Labor force/TA (LF)	No. of Employees(29)/Total Assets	3
Composition of Assets		
1) Inventory/TA (INV)	Inventory (78 or 3)/Total Assets	1,3
2) Accounts Receivable/TA (AR)	Accounts Receivable (2)/Total Assets	3
3) Capital Expenditure/TA (FCE)	Firm Cap. Exp.(30)/Total Assets	1,3
4)Provision for Doubtful Rec. (PDR)	Doubtful Rec. (67)/Total Rec. (67+2)	3
Cost/Performance Measures		
1) Change in Earnings (CE)	(Inc. before Ext. Items(18)- 48)t - (18-48)t-1	4
2) Earnings/Total Assets (ETA)	(18-Ext. Items and Disc. Oper. (48))t /Total Assets	1
3) Gross Margin % (GMP)	(Sales (12) - Cost of Goods Sold (41))/Sales	2,3
4) Sales & Admin. Exp./TA (SAE)	S&A(132)/Total Assets	3
5) % change in Sales (PCS)	(Sales t-Sales t-1)/Sales t-1	2
6) Dep./PPE (DPPE)	Dep.(14)/((PPE t (8)+PPE t-1)/2)	2
7) Research & Development/TA (RD)	R&D(46)/Total Assets	3
8) Oper. Inc./TA (OI)	Operating Inc. (13-14)/Total Assets	1
Other		
1) Order Backlog/TA (OB)	Order Backlog (98)/Total Assets	3
2) Div. Per Share/Price per Share (DSPS)	DPS t (26)/Close Price(24)	1,2
3) Repayment of LTD/LTD (RLTD)	Rep. Of LTD(114)/((LTDt(9)+LTDt-1)/2)	1
4) Change in Effective Tax Rate (CETR)	ETR t - ETR t-1*	3
5) LIFO Earnings (LIFO)	0 for LIFO, 1 for FIFO (59)	3
6) Audit Qualification (AQ)	0 for Unqualified, 1 for Qualified (149)	3
*ETR = 63/(18+63+49-48-55)		
Operating Income = Operating income before dep. (13) - Depreciation (14)		
•Variables from 1) Ou and Penman 2)Holthausen and Larcker 3) Lev and Thiagarajan		
•4) Variables suggested as risk proxies		

( ) Shows the Compustat data item number

3.2 Dependent Variables

The dependent variable is the excess returns calculated for a twelve-month period, beginning from the fourth month after the fiscal year end to allow for all of the firms to issue their financial statements. The return measures used in this study are market-adjusted and size-adjusted returns. Since the study is interested in excess market returns, market-adjusted returns may be the appropriate measure. However, prior research shows that size influences much of the market data and must be controlled. One way to control for size is to use size adjusted returns. The use of size-adjusted returns along with the inclusion of various measures of size as candidate variables should adequately control for firm size in this study (Greig, 1992; Holthausen and Larcker, 1992). Greig (1992) suggests that size explains a significant portion of the market-adjusted returns using the approach of Ou and Penman (1989). The market-adjusted returns are calculated as follows:

$$MAR_{im} = \sum_{t=1}^{m} (1 + R_{it}) - \sum_{t=1}^{m} (1 + R_{Mt}) \qquad (1)$$

where R_{Mt} is the return on the market in period t and R_{it} is the monthly return on asset i in period t from CRSP, $t = 1$ at the fourth month after a year-end, time m will be a 12-month holding period. The size-adjusted returns are calculated as follows:

$$SAR_{im} = \sum_{t=1}^{m} (1 + R_{it}) - \sum_{t=1}^{m} (1 - R_{st}) \qquad (2)$$

where R_{st} is the value-weighted return of the appropriate size portfolio and R_{it} is defined above. The size portfolios are constructed as in Holthausen and Larcker (1992) using CRSP to determine the appropriate size portfolio. The returns on the value-weighted size portfolios are then calculated from CRSP. The excess returns are set up as a binary variable, Y_i. For positive excess returns $Y_i = 1$, and for negative excess returns $Y_i = 0$.

The output of the logistic regressions are used to determine the probability that an individual firm will have excess returns for the 12-month period starting with the fourth month after the firm's year-end to allow for the financial statement information to be assimilated into the market. The logit model estimates the probability p_i :

$$\begin{aligned} logit(p_i) &= \log\left[\frac{p_i}{1 - p_i}\right] \\ &= B_0 + B_1 X_{i1} + B_2 X_{i2} + \ldots + B_{nX_\epsilon} \qquad (3) \\ p_i &= \frac{1}{1 + e^{-BX_i}} \end{aligned}$$

where $p_i = P(Y_i{=}1)$, and X_{i1}, ..., X_{in} are the candidate variables from table one. The B_j's are estimated by maximizing the log-likelihood function:

$$\sum_{i=1}^{N} L(B\,Y_i) = \sum_i \left(\frac{1}{1+e^{-BX_i}} \right) \qquad (4)$$

where B is the vector of parameters to be estimated, N is the number of observations, and Y_i is the value for each observation. Maximization leads to a series of nonlinear equations that are solved for the B_j's by iterative mechanisms. See Myers (1990) for these models.

3.3 Statistical Models

The study develops a logistic regression model for each industry listed in table 2 and for all of the industries combined. The models use five-year periods based on the returns from the twelve-month period beginning the fourth month after the firms' year ends. The logistic regression for a particular industry generates a score representing a probability of each firm in that industry classification having positive or negative excess returns. The time period 1978-1987 is used for model development to allow for comparison with the prior studies. The models developed from the industry regressions are then used to calculate the probability that the next period will generate positive or negative excess returns for each firm in that particular industry. The study combines all firms' probability scores. Firms in the top quartile are purchased long and the firms in the bottom quartile are sold short. This causes a difference in the total investment so weighted returns are used for each portfolio to allow for an equal investment position.

Table 2: List of Industries

# of firms in Compustat PC Plus	2 Digit SIC code	Category
308	13	Oil and Gas Extraction
165	20	Food and Kindred Products
463	28	Chemicals and Allied Products
519	35	Machinery and Computer Equipment
489	36	Electronic and Electrical Equipment
165	37	Transportation Equipment
480	38	Measurement Instruments
228	48	Communications
330	49	Electric, Gas and Sanitary Services
223	50	Durable Goods – Wholesale
209	60	Depository Institutions
152	61	Non Depository Credit Institutions
151	63	Insurance Carriers
295	67	Holding and Other Investment Offices
479	73	Business Services
184	80	Health Services
151	87	Engineering and Management Related Services

Stepwise logistic regression is used to select the variables that enter each of the logit models. Each of the seventeen industries used is listed in table 2 along with the number of firms in each industry. Regression models are estimated for two five-year periods 1978 to 1982 and 1983 to 1987. The regression model for 1978 to 1982 is used to predict returns for the five-year period 1983 to 1987, and the 1983 to 1987 model is used for predictive purposes for the five-year period 1988 to 1992.

3.4 Hypothesis Development

The most appropriate modeling technique for complex nonlinear phenomena may not be logistic regression (Zapranis and Refenes, 1994). It is possible that market phenomena are complex nonlinear phenomena as illustrated by the prior review. One of the most prevalent modeling techniques for this type of phenomena is a neural network. Therefore, neural networks are used to develop models to predict excess returns. If this technique is more successful, then the returns generated from these models should be greater than those generated with the logistic regression models. This leads to the first hypothesis:

> H_1: The returns from the neural network models will be less than or equal to the returns from the market-wide logistic regression models.
>
> H_{1a}: The returns from the neural network models will be greater than the returns from the market-wide logistic regression models.

Genetic algorithms are another complex modeling technique that is gaining wider acceptance in financial market research. Genetic algorithms address some problems associated with neural networks. One problem is the time required to develop an appropriate network. Another problem is the lack of explanatory capabilities in a neural network. A dual-layered genetic algorithm approach should reduce the time required to develop an appropriate model and allow for good explanatory capabilities (Eberhart, 1992) while still capturing the market complexity. This leads to the next hypothesis to be tested.

> H_2: The returns from the dual-layered genetic algorithm models will be less than or equal to the returns from the market-wide logistic regression models.
>
> H_{2a}: The returns from the dual-layered genetic algorithm models will be greater than the returns from the market-wide logistic regression models.

Neural networks and the dual-layered genetic algorithm approach are discussed in the next chapter. Hypothesis H_1 and H_2 are tested with paired t-tests. The mean of the returns U_{lr} and U_{nn} and U_{lr} and U_{ga} are tested using a one-tailed t-test to determine if the returns from the neural network and dual-layered genetic algorithm models are significantly greater than the returns from the market-wide logistic regression model. U_{lr} is the average of the returns from the market-wide logistic regression models. U_{nn} is the average of the returns from the neural network models. U_{ga} is the average of the returns from the dual-layered genetic algorithm. These averages will be from the 12-month holding period for all ten years of logistic regression, neural network, and genetic algorithm models.

Another area of interest is the possible effect of industry on the models. Industries are not consistent in their use of the variables selected for this study. For example the business service firms would be unlikely to have significant inventory and capital expenditure, while manufacturing firms would require proper management of these items. Research and development expenditures similarly have a greater importance in communications firms than depository institutions. Order backlog may be important in the transportation equipment industry as few orders are placed for commercial aircraft Maintaining a future output stream is critical to the transportation equipment industry while the insurance carrier industry should not be concerned with the variable. These examples of differing variable importance by industry illustrate the potential for better modeling on an industry-specific basis. Due to the number of observations needed to produce an effective model, only industries with a minimum of 150 firms in 1992 Compustat PC Plus are used in this study. (See table 2). This allows for removing firms with missing data points.

Logistic regression models are developed for each industry. These industry-specific models are used to obtain the probability that the firm has excess returns in the next five-year period. All firms are then combined and ranked by the probability of excess returns determined by the appropriate industry-specific model. The returns are value weighted to allow for equal investments in each portfolio. To obtain the market-wide model, the study combines all of the firms used in the various industries and creates a logistic regression model for the set of firms as a whole. The firms are then grouped in quartiles. The firms with the highest probability of having excess returns in the next twelve-month period are in the first quartile and the firms with the lowest probability of having excess returns are in the fourth quartile. The trading strategy requires the long purchase of the first quartile and the short sell of the fourth quartile firms. The same trading strategy and measures of excess returns as listed for the industry-specific models are used in the market-wide model. If the environment is different by industry, then the excess returns from industry-specific models will be greater than the excess returns from using a market-wide model. This leads to the last hypothesis tested.

H_3: The returns from the industry-specific models will be less than or equal to the returns from the market-wide logistic regression models.

H_{3a}: The returns from the industry-specific models will be greater than the returns from the market-wide logistic regression models.

A paired one-tailed t-test, using the average excess returns for the industry-specific models, U_{ilr}, and the market-wide models, U_{lr}, for each year from 1983 to 1992 tests this hypothesis. Table 3 lists a summary of the various returns needed to test each hypothesis.

Table 3: Logistic Regression Models

Hypothesis	Industry Level		
	Market	Individual	
H_{1a}: The returns from the neural network models will be greater than the returns from the logistic regression.			
H_1	T		$U_{nn} > U_{lr}$
H_{2a}: The returns from the genetic algorithm will be greater than the returns from the logistic regression.			
H_2	T		$U_{ga} > U_{lr}$
H_{3a}: Industry-specific models will generate greater excess returns than using a market-wide model.			
H_3	T	T	$U_{ilr} > U_{lr}$

3.5 DATA

Standard and Poor's compiles financial information for firms and stores twenty years of information in Compustat files. These files are the source of all financial data used in the study. The files are stored in computerized form that allows retrieval by industry. The Center for Research in Security Prices at the University of Chicago (CRSP) contains the historical price information used in this study. The two sets of files are linked by using the 6-digit CNUM and the first 2-digits of CIC from Compustat with the 8-digit CUSIP from the daily CRSP files. For each industry four models are developed: size-adjusted and market-adjusted returns using neural network and genetic algorithm methodologies. The first model developed corresponds to the Holthausen and Larcker study. Five-year models are constructed from 1978 to 1982 and 1983 to 1987. The models are used in the trading strategy for the following five-year periods from 1983 to 1987 and 1988 to 1992. The industries selected are those 2 digit SIC codes with more than 150 firms on the 1992 Compustat PC PLUS CD-ROM. See table 2 for a listing of the SIC codes used. The market adjusted and size adjusted returns are retrieved from CRSP. The data for the firms were collected from Compustat using the ARES9408, APST9408 and AFCV9408 files.

The PERMNO for each of the firms with the market value variable from Compustat is used to extract the daily compounded annual returns from the AMEX/NYSE or NASDAQ daily return files. The size-adjusted return was calculated by getting the daily compounded annual return for each firm for which the size decile was included on the CRSP file at the beginning of the year. The market valuation for each firm was also retrieved from CRSP for each firm. The appropriate decile value-weighted return is then calculated from these firms.

$$d_i \ CAR = \frac{\sum_{i=1}^{j} (MV * CAR)}{\sum_{i=1}^{j} MV} \tag{5}$$

Where MV is the end of year market value and CAR is the cumulative abnormal return variable form the CRSP file. CAR is compounded over each annual period.

This chapter reviewed the basic models tested in the study. The first part of the models discussed was the selection of the independent variables. The next component of the models discussed was the definition of the measures for the dependent variables. The time-periods these models were designed and tested in were indicated. The hypotheses regarding these models were developed and formally stated. The final item discussed was a review of the basic models' data sources. The next chapter expands on the discussion of the study's complex nonlinear algorithms.

CHAPTER 4
COMPLEX NONLINEAR ALGORITHMS

4.0 Introduction

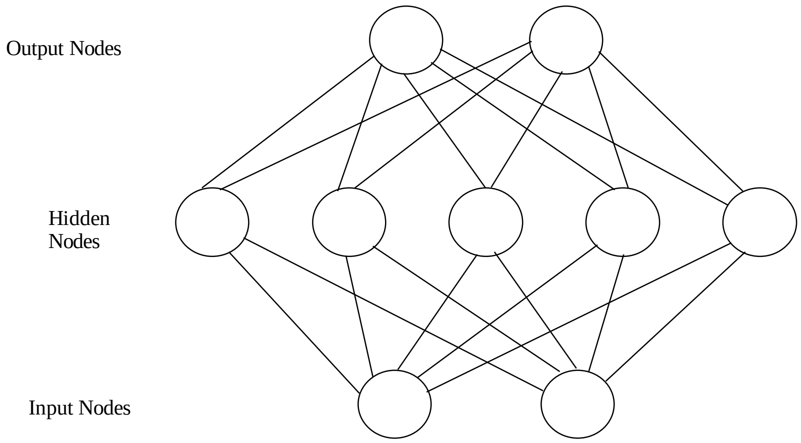
This chapter discusses the two complex nonlinear modeling techniques used in this study: neural networks and dual-layered genetic algorithms. The success of investment strategies based on these techniques is compared with strategies based on logistic regression.

4.1 Neural Networks

A Neural Network uses a network of nodes and connecting weights to represent the interaction between input and output parameters in a prediction model. The primary function of a neural network model is to assign appropriate weights to the input nodes of a network so that a weighted function of the input nodes predicts the outputs. Several algorithms can be used to develop these weights including feedforward and backpropagation. In financial market studies backpropagation is frequently used and therefore this study uses a backpropagation algorithm. The backpropagation algorithm calculates the error at the output nodes and passes it back through the hidden nodes to the input nodes adjusting the weights as the algorithm passes the error back to the input nodes.

This study used a three-layer, fully connected initial network. (See figure 2.) In this type of network the first layer is composed of the input variables, the second layer is composed of hidden nodes and the last layer is composed of the output nodes.

Figure 2 - Network Configuration



Four-layer, fully connected networks were also tested. However, the three layer networks performed as well as the four layer models with reduced complexity. The total input for any node *j* in the hidden or output layer is given by:

$$X_j = \sum_{i=1}^{l} W_{ij} X_i , \qquad (6)$$

where $i = 1, ..., l$ are all of the nodes in the previous layer and W_{ij} is the connection weight. The input X_j is then transformed using a nonlinear activation function producing a standardized Y_j for node *j*. The function used in this study was the sigmoidal function. The tanh function was also tested but prediction accuracy was highest with the sigmoidal function. The sigmoidal function has a functional form equivalent to the logit function:

$$Y_j = \frac{1}{1 + e^{-X_j}} \qquad (7)$$

Removing the hidden layer from the neural network and leaving one output node in the final layer, the output Y_j generates a probability estimate p_j similar to the logit model. However, the weights W_{ij} in the neural network and the B_js in the logit model are not equivalent according to Sen et al. (1994). The weights W_{ij} are computed by minimizing the error in estimating the output, while maximizing the log-likelihood function computes the B_js. The neural network uses a pattern matching process while the logit model tries to estimate a partition of sample space.

The neural network has an additional layer of hidden nodes that allow it to generate complex nonlinear mappings that may be useful in predicting chaotic phenomena such as stock market returns (Zapranis and Refenes, 1994; Kuan and White, 1994). Logit is a single layer approach that cannot develop such complex nonlinear mappings. A predetermined functional form that fits the data to the model constrains the logistic regression, while no predetermined functional forms constrain the neural network. Therefore, the neural network can produce any functional form necessary to map complex phenomena (Bansal et al., 1993; Malliaris, 1994; Dutta and Shekhar, 1988).

In this study the initial weight matrix was generated by random numbers between .-1 and +.1. The weight matrix is updated after each training set is reviewed by minimizing the error. The error function is:

$$\mathrm{E} = \frac{1}{2}\sum_{k}(y_j - d_j)^2 \qquad (8)$$

The function above is minimized to obtain the weights W_{ij}, where *j* is an index of nodes in the output layer and *k* is the set of nodes in the output layer, y_j is the computed output and D_j is the desired output. The error is averaged for the entire training set and the change in weight)W is passed back through the edges connecting the hidden layer to the output layer and the input layer to the hidden layer. This is done by:

$$\Delta w \; = \; E * Y_j (1 - Y_j) Y_i \qquad (9)$$

Once the weights are updated, the training set is run through the updated model to determine the predictive accuracy. If a value greater than .5 is obtained for the output node then positive excess returns are expected; a value less than .5 indicates negative excess returns are expected. A test set is run through the same weight matrix and the number of examples that it predicts with positive excess returns or negative excess returns is computed. The training and updating of weights continues until the test sample performance deteriorates.

A sensitivity analysis is performed when the test sample performance deteriorates. Sensitivity analysis is a method of determining the impact inputs have on the outputs of a network. Each input is changed slightly to determine the effect on the output. The program used noted the effect on the output in a percentage format. The total of all of the input effects is equal to 100 percent. This allows for the selection of an input node for pruning when it has little or no effect on the output node.

This analysis is used to prune unimportant nodes. The importance of the nodes is based on the impact on the output nodes. If the sum of W_is is small then the node will have little impact on the classification (Giles andOmlin, 1994) and will be pruned or removed. Before removing any input variables due to little impact, three models were developed using the most appropriate network structure. The average influence on the output of each input variable was determined. All input nodes had some impact on the desired output. Therefore no input nodes were determined to be unimportant and pruned. Pruning makes the neural network model more generalizable and helps avoid over fitting (Reed, 1993). The holdout data was then run through the model and the highest 25 percent of the scores for the output node were bought long and the lowest 25 percent of the scores were sold short. The calculation of the return for each portfolio is value weighted to allow for an equal initial investment. The number of nodes used was restricted based on the number of training patterns available. A guideline used for the maximum number of nodes is based on (LeFebvre, 1995):

$$N_i * N_h * N_o \leq \frac{TP}{PL} \tag{10}$$

where N_i is the number of input nodes, N_h is the number of hidden nodes, N_o is the number of output nodes, TP is the number of training patterns and PL is the precision level desired.

4.2 Dual-Layered Genetic Algorithms

The dual-layered genetic algorithm approach combines the use of a recursive-partitioning algorithm with genetic algorithms. The recursive-partitioning algorithm develops a set of if-then rules used to classify all firms in the appropriate sample based on having positive or negative excess returns. The algorithm used was based on the Automatic Interaction Detection (AID) process, first proposed by Morgan and Sonquest (1963). AID uses an interval scaled dependent variable to maximize the between-group-sum-of-squares (F-statistic) to select the variables in the model. Each rule set predicts excess returns. However, this study uses the rule sets as the starting point for the genetic algorithms.

Genetic algorithms can be used as classifier systems consisting of a set of rules (Holland, 1992). A genetic algorithm operates on strings of bits that contain information often called chromosomes or genotypes. Each bit of the string represents the presence or absence of a given characteristic. A complete string represents a combination of the presence or absence of all the relevant characteristics for a particular problem. This string contains a strategy to solve the problem. A population of strings is generated and evaluated by mapping them to a solution to the task at hand. This evaluation determines the performance of the strategy represented by each string. The best strings are modified in the search for improved strategies. These modifications include mating two strings through crossover. In figure 3 there is an example of crossover. Two strings composed of five variables serve as parents. The strings are randomly split between the second and third variables. The code for the first parent's beginning two variables are combined with the second parent's remaining variables, and the second parent's first two variables are combined with the first parent's remaining variables. Another modification used is mutation, where individual bits are randomly switched. The third section of figure 3 shows an example of this. The third bit in the first child is randomly selected and switched from a 1 to a 0, and the fifth bit in the fourth variable for the second child is randomly switched from a 0 to 1. Genetic algorithms use these string modifications, or operators to find the best string or strategy (or disjunct) to explain a problem.

Figure 3 - Genetic Operations

Parents:

Var. 1 (31)	Var. 2 (15)	Var. 3 (8)	Var. 4 (4)	Var. 5 (24)
1 1 1 1 1	0 1 1 1 1	0 1 0 0 0	0 0 1 0 0	1 1 0 0 0

Var. 1 (13)	Var. 2 (19)	Var. 3 (3)	Var. 4 (22)	Var. 5 (5)
0 1 1 0 1	1 0 0 1 1	0 0 0 1 1	1 0 1 1 0	0 0 1 0 1

Offspring via Crossover

Var. 1 (31)	Var. 2 (15)	Var. 3 (3)	Var. 4 (22)	Var. 5 (5)
1 1 1 1 1	0 1 1 1 1	0 0 0 1 1	1 0 1 1 0	0 0 1 0 1

Var. 1 (13)	Var. 2 (19)	Var. 3 (8)	Var. 4 (4)	Var. 5 (24)
0 1 1 0 1	1 0 0 1 1	0 1 0 0 0	0 0 1 0 0	1 1 0 0 0

Mutation of Offspring (* Mutation)

Var. 1 (31)	Var. 2 (15)	Var. 3 (3)	Var. 4 (22)	Var. 5 (5)
1 1 1 1 1	0 1 *0 1 1	0 0 0 1 1	1 0 1 1 0	0 0 1 0 1

Var. 1 (13)	Var. 2 (19)	Var. 3 (8)	Var. 4 (4)	Var. 5 (24)
0 1 1 0 1	1 0 0 1 1	0 1 0 0 0	0 0 1 0 *	1 1 0 0 0

Another operator that needs to be selected is the crossover type. The crossover types available for modeling are one-point, where strings are randomly divided at the same location and then one segment of the strings are switched, and two-point where strings are randomly divided at two locations and one section of each string is switched and uniform where each individual bit can be passed on from either string. The uniform crossover can find unnoticed combinations but it can also lose positive combinations, causing a significant drop in performance (Davis, 1991). The initial strings are generated from the decision tree rules, and the initial positive combinations should be retained. Therefore, a one-point crossover type was used. The crossover probability determines how many strings in the next generation are offspring produced through mating. A crossover probability of 0.6 was used based on Sikora and Shaw (1994). The last operator used was mutation. Mutation is used to help avoid local minimum and maximum solutions. Based on Sikora and Shaw (1994) only 1 in 100 bits weremutated .

Since the genetic algorithm begins with the disjunct from the recursive-partitioning algorithm and always retains the best disjunct in each generation, the final disjunct given by the genetic algorithm is at least as good as the best disjunct it started with (Sikora and Shaw, 1994).

Because of the learning involved with the genetic algorithm through the generations, the final disjunct's yield should be better than the best rule set provided by the recursive-partitioning algorithm.

The first step in this part of the study was to use the candidate variables in table 1 and the binary excess return variable in the recursive-partitioning algorithm. Kass (1980) refined the AID process by using a nominal scaled dependent variable, and maximizing the significance of a chi-squared statistic at each partition. The final model, or rule set, is developed using significance testing and an exhaustive search algorithm for the highest chi-squared grouping within each variable. This process continues until no significant splits are found. There may be more than one significant split for each partition. The three best partitions for predicting both excess and non-excess returns as determined by classification accuracy with a reasonable number of firms were used for each sample. This study used KnowledgeSeeker, a commercial software package based on Kass's algorithm (Biggs, et al., 1991). The model works with ordinal and nominal data and has no limitations imposed by non-continuous or non-normal data.

Each training set included in the market-wide logistic regression was used to develop rule sets with KnowledgeSeeker's recursive partitioning. The genetic algorithm initially generated 50 random strings of up to 5 variable combinations. The worst six initial strings were removed and replaced by the best three partitions for predicting excess returns and the best three partitions for predicting non-excess returns from KnowledgeSeeker before the next generation was developed. Sikora and Shaw (1994) used an initial population size of 42 strings for a bankruptcy analysis and 46 for a loan classification problem, therefore fifty strings appears to be an appropriate starting population.

The data were converted into integers from 1 to 10 based on the decile of each value. The genetic algorithm created a range based on deciles for each of the variables except LIFO and AQ, which were dummy variables.

The fitness function added 1 if an example matched all of the given variables and the excess value, resulting in a positive example, and a -1 if all the given variables matched but the excess value did not match, resulting in a negative example. The genetic algorithm used a random sample of the firms in each testing population. The testing populations were composed of 500 positive and 500 negative examples. This was done to promote the existence of disjunct or strings that can be used to predict both positive and negative returns. For this study both were equally important. No combinations of more than 5 variables produced a high fitness score since few firms have the same decile in more than 5 variables. The genetic codes generated were applied to each testing period in descending fitness score until approximately 25 percent of the firms were selected to buy long or sell short. The portfolio return was value weighted to allow for an equal investment.

There are two principal methods of deriving rules using a genetic algorithm. One is the Pitt approach where each example represents a complete concept. In this approach each string is of variable length and represents a complete solution. Another method uses the Michigan approach where each member of the population may be a single disjunct. Each disjunct is of fixed length but only provides a solution to part of the sample space. The separate disjuncts are then combined to obtain the complete solution (Janikow, 1993; Sikora and Shaw, 1994). Sikora and Shaw use the Michigan approach in the initial testing of this dual-layered methodology. They cited two reasons for using the Michigan approach "first, it does not demand an a priori assumption of knowledge of the number of disjuncts, and second, it lets the individual disjunct compete against each other, thereby allowing for a more powerful disjunct to be discovered." Since this study was looking for the most concise rule set and had no prior conception of the number of disjuncts needed, this was considered the best approach for the current study.

At each generation the genetic algorithm retains the best disjunct and replaces the rest through genetic operations. Once the genetic algorithm converges, the best hypotheses found are retained, and the positive examples it covers removed. In this study positive examples were those in which the hypothesis correctly predicted positive excess returns. Convergence occurs when no further improvement in the fitness measure occurs and the population reaches stability. The process is repeated until all the positive instances are covered. The final rule is a combination of all the hypotheses found. This procedure for searching the instance space (I-space) is called explanation-based filtering (Sikora and Shaw, 1994).

The investing strategy used for this section was rule based. The final rule from the genetic algorithm contained many disjuncts, covering the appropriate firms in the sample. For the trading strategy the disjuncts were applied separately. The first rule was from the first hypothesis, or disjunct retained, then a second rule was generated from the combination of the first two hypotheses retained. The rules were applied individually in the test period until they selected approximately 25 percent of the firms.

The same approach was used to select the firms for the sell short group. The model was developed to select those firms with a negative excess return in the training and testing periods. The firms were selected until the initial investment equaled the amount invested in long positions. The firms added by the last rule were randomly added to equal the long position investment.

The success of the genetic algorithm depends on the proper selection of various operators. The most important operator is the fitness function used to select those strings to be retained. The fitness function chosen for this classification problem is based on the need to maximize the difference between excess and non-excess firms. The most straightforward fitness function is:

$$F_c(C) = p - n \qquad (11)$$

where p and n are the number of positive and negative examples covered by the condition$c$. The fitness function depends on the specific application for which the genetic algorithm is being used. A major objective of the study was to develop a trading strategy using 25 percent of the total firms as buy and hold and 25 percent of the total firms as sell short. This fitness function allowed for the development of disjuncts that predicted excess returns and non-excess returns. A high absolute fitness score indicated a good discriminating string was found. If the fitness score was negative then changing the value of the excess term retained the disjunct.

The objective was to buy only firms that generated positive excess returns. A firm with negative excess returns reduced the overall profitability of the trading strategy and should have been avoided. Therefore, the penalty term was required. In Sikora and Shaw (1994) the penalty term of (1/n) for covering negative examples was used. In this study covering a negative example was just as important as covering a positive example. The simple fitness score of positive examples covered, less negative examples covered, gives equal weighting to both examples.

This chapter discussed the complex nonlinear modeling techniques used in this study:

neural networks and dual-layered genetic algorithms. The neural network discussion reviewed the general components of neural networks and the selection of the specific components in this studies' neural networks. The dual-layered genetic algorithm methodology was discussed in general along with the parameters used in the current study.

CHAPTER 5
RESULTS

5.0 Introduction

The results are reported in three major sections. The first section shows the summarydata and describes the market-wide logistic regression models developed. The returns generated by the trading strategy based on the market-wide logistic regression models are reported. These returns are compared with the returns from Holthausen and Larcker's and Ou and Penman's studies. The next section discusses the results of the neural network models and statistical tests if the returns from the neural network models are greater than the returns from the market-wide logistic regression models. The third section reports the results from the dual-layered genetic algorithm model and tests if the returns from the dual-layered genetic algorithm model are greater than the returns from the market-wide logistic regression model. The last section reviews the industry-specific models and tests if the industry-specific logistic regression models generated greater returns than the market-wide logistic regression models.

5.1 Market-wide Logistic Regression

The first major procedure in this study was to develop market-wide logistic regression models. An important aspect of the development of the models was to locate and account for outliers. Tukey=s interquartile range with outer fences (Hildebrand and Ott, 1987) locates outliers for normal data. A univariate normality test was performed on each variable to determine if Tukey's method was appropriate. The only variables that did not violate the Martinez-Iglewicz normality test (Hintze, 1995) were MV and ATA. The normality assumption was rejected for both of these variables using the Kolmogorov-Smirnov test (Hintze, 1995). The summary information for the data is shown in table 4. The data for the model building periods 1 and 2 do not include outliers.

Table 4 - Variable Summary Data

	Period 1			Period 2			Period 3		
	Mean	Standard Deviation	Count	Mean	Standard Deviation	Count	Mean	Standard Deviation	Count
MV	3.8933	1.9104	11,357	4.0175	2.0294	14,380	4.3334	2.1858	15,980
ATA	4.4233	2.3866	11,356	4.2794	2.5045	14,381	4.7184	2.5135	15,980
BVMV	1.5824	1.7569	11,190	1.2182	1.7937	14,196	1.3399	1.9624	15,460
LF	0.0192	0.0258	10,425	0.0154	0.039	11,773	0.0103	0.0161	14,044
INV	0.1138	0.1015	4,758	0.0969	0.0936	6,343	0.094	0.0912	6,863
AR	0.2417	0.1589	10,674	0.2264	0.1628	13,268	0.2346	0.1847	15,130
FCE	0.1114	0.1102	9,712	0.0917	0.0959	12,416	0.0771	0.0896	13,269
PDR	0.09	0.2141	7,586	0.0985	0.1934	9,528	0.0752	0.1273	10,751
CE	2.4552	38.4498	11,346	2.6961	83.076	14,351	3.5838	209.8869	15,955
ETA	0.0281	0.1505	11,300	-0.032	0.3026	14,199	-0.0298	0.3911	15,910
GMP	0.3384	0.6577	11,226	0.1773	2.395	14,041	-0.9079	56.0048	15,299
SAE	0.2977	0.2364	2,395	0.3599	0.2882	5,712	0.3685	0.3714	9,024
PCS	0.1267	0.7158	11,230	0.0718	0.8698	14,004	-873.3554	43156.6307	15,311
DPPE	0.1566	0.1274	9,829	0.23	0.3763	12,646	3.0265	166.5172	14,006
RD	0.055	0.0824	4,927	0.09	0.1383	6,789	0.1042	0.2269	7,359
OI	0.0777	0.1764	10,969	0.0056	0.2944	13,811	0.0084	0.3561	14,827
OB	0.6473	0.7498	3,760	0.5469	0.6848	4,517	0.5489	0.8863	4,688
DSPS	0.0498	0.0351	6,898	0.0452	0.046	6,123	0.0551	0.1672	6,540
RLTD	0.412	0.9488	8,300	0.8828	5.7936	10,241	118.3475	11165.1059	11,822
CETR	-0.0201	0.1685	6,845	-0.0764	1.177	8,380	-186.1646	18250.4618	9,615

Tukey's interquartile range was modified, since the variables were not univariate normal. One common methodology to address non-normality is to use rank orderings; therefore a modification was developed based on the ordering of the variables. Tukey's method eliminates any observation greater than or less than the outer fences as outliers. Since the data were skewed and not normal the method was modified. The starting point of the modified methodology was Tukey's outer fence values. The second observation outside the outer fence was divided by the first observation outside the outer fence to measure the distance between the ordered observations. This process continued, moving further from the outer fence until the ratio was greater than 1.25. Once this ratio was exceeded the remaining observations were deemed outliers and removed from the model building periods. The cut-off levels for each variable in both periods are in appendix table A1.

The next step in developing the market-wide logistic regression model was to eliminate all firm years with missing variables. Removing these firm years led to the elimination of all but 135 firm years in period one and 186 firm years in period two. A logistic regression model was created for market-adjusted returns and size-adjusted returns as calculated from equations one and two respectively. Due to the small number of firm years left in themodel building period, additional models were developed by removing variables that eliminated large numbers of firm years.

A second model was developed after removing the variable Sales and Administration Expense/Total Assets (SAE), which had the most missing values (8,962) in period 1. After eliminating the SAE variable and removing the firm years with missing values there were 527 firm years in period one and 385 firm years in period two. A third model was developed after removing Order Backlog/Total Assets (OB), which had 7,597 missing values in period 1. These models had 978 firm years in period 1 and 766 firm years in period 2. The next model was developed after removing Inventory/Total Assets (INV), which had 6,599 missing values in period 1. These models had 1,296 firm years in period 1 and 1,026 firm years in period 2. The fifth model was developed after removing Research and Development/Total Assets (RD), which had 6,430 missing values in period 1. These models had 2,328 firm years in period 1 and 1,791 firm years in period 2. The final model was developed after removing Change in Effective Tax Rate (CETR), which had 4,512 missing variables in period 1. These models had 3,730 firm years in period 1 and 4,124 firm years in period 2. Three additional variables had missing values in 20 percent or more firm years in both periods, Provision for Doubtful Receivables (PDR), Dividends Per Share/Price per Share (DSPS) and Repayment of Long Term Debt/Long Term Debt (RLTD). However these variables had been determined significant in four or more of the previous eighteen models and were not removed. The complete listing of missing firm years is in appendix table A2.

Each model was tested against all of the firm years in the appropriate model-building period. Two accuracy measures were calculated: the total correct percentage of predictions for all firms and the correct percentage of predictions for excess return firms. Since the prior studies' results appeared to be driven by the short position, extra weighting was given to the predictive ability for the excess firms or those invested in the long position in an attempt to improve the results of the long position. The main criteria used to determine the most appropriate model was the highest average prediction accuracy of the two measures. One additional criteria was used in selecting the model: the total prediction accuracy rate had to be greater than 50 percent since the main objective was to generate total returns.

The final size-adjusted return model for period one excluded the candidate variables: SAE, OB, INV, RD, and CETR. The regression models for each of the 6 data sets and the associated accuracy rates are in appendix table A3. The final regression model was:

P_r = -0.9406 + 0.0783 * ATA + 2.6064 * LF - 0.9642 * FCE + 2.1330 * PDR + 1.8504 * OI.

This model included two size variables Average Total Assets (log) (ATA) and Labor Force/Total Assets (LF), two composition of assets measures Capital Expenditure/Total Assets (FCE) and Provision for Doubtful Receivables (PDR), one cost/performance measure Operating Income/Total Assets (OI), and no other variables.

The final market-adjusted return model for period one excluded the candidate variables: SAE, OB, INV, and RD. The regression models for each of the 6 data sets and the associated accuracy rates are in appendix table A4. This final model was:

P_r = -0.2887 + 0.6710 * AR - 0.4877 * PCS + 1.4402 * OI - 0.1927 * RLTD.

This model included no size variables, one composition of assets measure, Accounts Receivable/Total Assets (AR), two cost/performance measures, Percentage Change in Sales (PCS) and OI, and one other variable, Repayment of Long Term Debt/Long Term Debt, (RLTD).

In period two the final model for both size-adjusted returns and market-adjusted returns excluded the candidate variables: SAE, OB, INV, RD, and CETR. The regression models for each of the 6 data sets and the associated accuracy rates are in appendix tables A5 and A6. The final size-adjusted model was:

P_r = -0.6564 - 0.2808 * MV + 0.3298 * ATA - 0.0761 * BVMV - 1.0747 * FCE + 2.1468 * OI + 0.2299 * LIFO.

This model included three size variables Market Value (log) (MV), ATA and Book Value/Market Value (BVMV), one composition of assets measure, FCE, one cost/performance measure, OI, and one other variable, LIFO Earnings (LIFO). The final market-adjusted model was:

Pr = -0.9930 - 0.1324 * MV + 0.2236 * ATA - 1.0417 * FCE + 1.8618 * OI + 0.1790 * LIFO.

This model included two size variables, MV and ATA, one composition of assets measure, FCE, one cost/performance measure, OI, and one other variable, LIFO.

These models were then used to determine which firms to purchase long or sell short for the following five-year testing periods. Because the logistic regression indicated these variables were significant, only those firms in the trading period that had all of the variables in the appropriate model were used. The return generated from purchasing 25 percent of the firms with the highest probability of generating excess returns and short-selling 25 percent of the firms with the lowest probability of generating excess returns averaged .140444 for market adjusted returns over the 10-year period (Table 5). The size-adjusted average return over the time period was .175603 (Table 6). The detailed returns by year and position are in appendix tables A13 and A14.

Table 5 - Market Adjusted Excess Returns

	Time Period	Number of Firms in Short Portfolio	Short Return	Number of Firms in Long Portfolio	Long Return	Total Return	
	1983-1992	5,517	-0.129537	5,517	0.010907	0.140444	value-weighted
H & L	1978-1988	12,289	-0.042300	11,774	0.000300	0.042600	equal-weighted
O & P	1973-1983	1,829		8,119		0.083400	value-weighted

Note: Ou and Penman's reported results (Table 6) do not include the value of each portfolio so no comparison can be made regarding the exact amount of the total return due to each investment position. However using an equal-weighted portfolio approach the long position would have a return of .035650 and the short position would have a return of -.058850.

Table 6 - Size Adjusted Excess Returns

	Time Period	Number of Firms in Short Portfolio	Short Return	Number of Firms in Long Portfolio	Long Return	Total Return	
	1983-1992	5,096	-0.153411	5,096	0.022192	0.175603	value-weighted
H & L	1978-1988	10,457	-0.047200	10,824	0.032500	0.079700	equal-weighted
O & P	1973-1983	1,814		8,118		0.055400	value-weighted
A & B	1974-1988					0.132000	

Note: Ou and Penman's reported results (Table 8) do not include the value of each portfolio so no comparison can be made regarding the exact amount of the total return due to each investment position. However using an equal-weighted portfolio approach the long position would have a return of - .00023 and the short position would have a return of -.06513. Abarbanell and Bushee (1998) (A & B) report the returns of a fundamental trading strategy based on Lev and Thiagarajan's signals. A & B only reported the results for size-adjusted returns.

There are two major studies that used a logistic regression methodology to generate excess returns. A comparison of the results of the current study with these prior studies should indicate the success of this current approach. Tables 5 and 6 compare the excess returns generated from each trading strategy. Tables 7 and 8 compare the prediction accuracy rates of the current study with L&T and O&P. The O&P accuracy rate is not fully compatible since they predict a change in earnings increase.

Table 7 - Market Adjusted Prediction of Excess Earnings or Returns

	Time Period	Number of Firms in Short Portfolio	Accuracy Rate	Number of Firms in Long Portfolio	Accuracy Rate	Total Accuracy
	1983-1992	5,517	69.6%	5,517	46.4%	58.0%
H & L	1978-1988	12,289	61.6%	11,774	42.1%	52.1%
O & P	1973-1983	1,829	70.0%	8,119	65.5%	66.3%

Note: Ou and Penman's accuracy rate is based on correctly predicting an earnings increase and not an excess return.

Table 8 - Size Adjusted Prediction of Excess Earnings or Returns

	Time Period	Number of Firms in Short Portfolio	Accuracy Rate	Number of Firms in Long Portfolio	Accuracy Rate	Total Accuracy
	1983-1992	5,096	70.3%	5,096	49.4%	59.9%
H & L	1978-1988	10,457	62.9%	10,824	45.9%	54.3%
O & P	1973-1983	1,814	70.0%	8,118	65.5%	66.3%

Note: Ou and Penman's accuracy rate is based on correctly predicting an earnings increase and not an excess return.

The current returns were greater than Ou and Penman's returns for the period from 1973 to 1983. Ou and Penman also used a value-weighted trading strategy to allow for an equal investment in both positions. Their reported results did not include the value of the firms used in each portfolio so no accurate breakdown by portfolio was possible. In an attempt to determine the breakdown of Ou and Penman's returns, an equal-weighted return was calculated from Ou and Penman's reported values[6] and the results are in tables 7 and 8. This indicated that for the market-adjusted return approximately 60 percent of the total excess return was due to the short position

[6] Ou and Penman values are shown in their tables 6 and 8.

and approximately 100 percent of the size-adjusted total excess return was due to the short position. This suggests that the short position dominated the results.

Using market-adjusted returns Ou and Penman showed a return of .0834, while size-adjusted returns earned a return of .0554. Ou and Penman selected the firms with a P_r value, or probability of excess earnings per share, with a value greater than .6 to invest in a long portfolio, and firms with a probability of excess earnings per share of less than .4 to invest in the short portfolio. Reviewing the number of firms in each position, it appears that the majority of firms were predicted to have excess earnings per share. The size-adjusted returns varied from 8,118 firm years in the long position to 1,814 firm years in the short position and the market-adjusted returns varied from 8,119 firm years in the long position to 1,829 firm years in the short position. This discrepancy in the probability values suggests that some improved model may generate greater excess returns. Holthausen and Larcker showed that using a binary variable for excess returns directly led to an improved model and greater returns from a similar trading strategy.

The returns of the market-wide logistic regression model were also greater than that of Holthausen and Larcker=s. The average size-adjusted return using the market-wide logistic regression model was .1756 compared to Holthausen and Larcker's return of .0797. The average market-adjusted return using the market-wide logistic regression model was .1404 compared to Holthausen and Larcker's return of .0426. Holthausen and Larcker used an equal-weighted return for the firms they selected using their Pr values. The use of equal-weighted returns may be part of the difference in the results. The value-weighted returns allow for a matching investment in portfolios and do not give extra weighting to small firms.

Reviewing Holthausen and Larcker's results indicate that approximately 100 percent of the market-adjusted returns and 60 percent of the size-adjusted returns were from the short position. This is further evidence that the short position drives the excess returns. Another difference between the current study and Holthausen and Larcker's study was the number of firms used in the trading strategy. Holthausen and Larcker invested long in the top and bottom 30 percent of the firms and used all industries. The current study used only the top and bottom 25 percent of firms in selected industries. Tables 5 and 6 show that Holthausen and Larcker used over twice the number of firms in each portfolio. If the P_r measure were a good indicator of excess returns then using more extreme tails would earn a greater excess return.

The total market would expect to have total excess returns of 0. However certain industries may have returns greater or less than 0. Since industries with fewer than 150 firms on the 1992 Compustat PC Plus CD-ROM were not used, the overall return of the sample selected may bias the results. As a check, all individual result tables indicate the total value weighted return for the entire trading population. This return was listed under the heading Population Return in appendix tables A13 to A23. The market-wide average should be zero, and no sets of firms had an average sample return less than -.02 percent. These values do not appear responsible for the excess returns found in this study. However, each set of firms' negative average sample return supports the statement that the short position was responsible for the excess returns.

Comparing the accuracy rates between the current study and Holthausen and Larcker=s study supports the greater returns generated in this study. The current study's accuracy rate in predicting non-excess returns in the short portfolio were 69.6 percent to 61.6 percent for Holthausen and Larcker in market-adjusted returns and 70.3 percent to 62.9 percent for size-adjusted returns. The accuracy rate for excess returns in the long portfolio was 46.4 percent to 42.1 percent for market-adjusted returns and 49.4 percent to 45.9 percent for size-adjusted returns. The current study had greater prediction accuracy in all portfolios.

5.2 Neural Network Returns

The next step was to determine the appropriate neural network models. For consistency purposes the training and trading period data used were the same firm year data as the logistic regression models. Initially, various configurations were attempted to map the data. The most appropriate model based on minimizing the square error was a configuration of one input layer with 18 nodes for the market return model for 1978 to 1982, and 17 nodes for the remaining models. The number of hidden nodes in the final model was nine for three models and eight for one model. One output node was used with a 1, 0 binary variable in all four models, 80 percent of the training period firm years were used to train the network and 20 percent were used as a holdout test. To avoid overtraining, training was stopped when the test sample error started to increase. At this point the network weights were saved and used to test the following five years.

A sensitivity analysis was performed and all of the input variables influenced the desired output.

These returns are significantly less than the returns using the market-wide logistic regression with an average return of .018228 for market adjusted returns and .082732 for size adjusted returns. See table 9 for the significance tests. The complete returns by year are listed in appendix tables A14 and A15.

Table 9 - T-test for H_1

	Number of Firms in Short Portfolio	Short Return	Number of Firms in Long Portfolio	Long Return	Total Return	Return Type
Market-Wide Logistic Regression	5,096	-0.153411	5,096	0.022192	0.175603	Size Adjusted
	5,517	-0.129537	5,517	0.010907	0.140444	Market Adjusted
Neural Network	5,096	-0.091094	5,096	-0.008362	0.082732	Size Adjusted
	5,517	-0.039123	5,517	-0.020895	0.018228	Market Adjusted
One tailed T-test value for Size Adjusted Returns					0.020181	
One tailed T-test value for Market Adjusted Returns					0.003064	

Using the actual return as the output node instead of a binary variable was an alternative that may have improved the neural network performance. However, changing the variable from a continuous variable to a binary variable may have caused the loss of information. The neural network architecture allows for the use of a continuous variable and therefore the actual returns were used in the output node. Neural networks were again trained in the same manner for each period and return, then applied to the appropriate testing period. The results of these tests are in appendix tables A20 and A21. These results indicate a return of -0.056290 for market-adjusted model and -0.006546 for the size-adjusted model. These returns are lower than those generated using the binary variable indicating that using a continuous variable captures more noise in the model.

If the neural network architecture is better suited for mapping complex market phenomena than logistic regression, the excess returns generated from basing a trading strategy on the neural network model should be greater than one based on the logistic regression model. The formal hypothesis is listed below.

H_1: The returns from the neural network models will be less than or equal to the returns from the market-wide logistic regression models.

H_{1a}: The returns from the neural network models will be greater than the returns from the market-wide logistic regression models.

To test if the returns are statistically different a one-tailed t-test was used. The results are shown in table 9. The results indicate that there is a significant difference in market adjusted returns at the .01 level and for size adjusted returns at the .02 level. These results indicate that the null hypothesis can not be rejected.

5.3 Dual-Layered Genetic Algorithms.

The dual-layered genetic algorithm approach requires the development of a set of initial classification rules using a rule induction algorithm. The data should be the same as that used in the development of the logistic regression and neural network models. However, genetic algorithms do not work with continuous data. Therefore, the data needed to be converted into a non-continuous form. The data were converted into decile form and KnowledgeSeeker used the converted data to develop the initial rule sets to predict the presence or absence of excess returns. The top three initial classification rules to predict the presence or absence of excess returns were converted into character strings and used as the genetic algorithms' first generation. The genetic algorithm used crossover and mutation parameters to arrive at the final disjuncts. The initial rules and final disjuncts and the associated fitness score are listed in appendix table A28.

The returns generated from this approach were significantly less than the returns from using the market-wide logistic regression. The average return was 0.037912 for market adjusted returns and 0.090556 for size adjusted returns. The results of the significance tests are in table 10. These results indicate that the null hypothesis can not be rejected. Appendix tables A16 and A17 list the complete returns by year. The dual-layered genetic algorithm approach should allow for a complete interpretation of the model. The formal hypothesis is stated below:

H_2: The returns from the dual-layered genetic algorithm models will be less than or equal to the returns from the market-wide logistic regression models.

H_{2a}: The returns from the dual-layered genetic algorithm models will be greater than the returns from the market-wide logistic regression models.

Table 10 - T-test for H_2

	Number of Firms in Short Portfolio	Short Return	Number of Firms in Long Portfolio	Long Return	Total Return	Return Type
Market-Wide Logistic Regression	5,096	-0.153411	5,096	0.022192	0.175603	Size Adjusted
	5,517	-0.129537	5,517	0.010907	0.140444	Market Adjusted
Genetic Algorithms	6,915	-0.084513	7,223	0.006043	0.090556	Size Adjusted
	8,025	-0.044507	5,098	-0.006595	0.037912	Market Adjusted
One tailed T-test value for Size Adjusted Returns					0.026892	
One tailed T-test value for Market Adjusted Returns					0.006022	

The rule for predicting excess market-adjusted returns for period 1 was: if the audit report was not qualified, the log of the market value wasn't in either extreme decile and the change in earnings was in the top 60 percent, then it was predicted as an excess return firm. The rule for predicting non-excess size-adjusted returns was: if the book value to market value ratio was not in the lowest decile and the labor force to total asset ratio was in the bottom 60 percent of firms then the firm was predicted as a non-excess return firm.

For size-adjusted returns in period 1 the final disjuncts for both excess and non-excess returns were based on the change in earnings. If the change in earnings was in the 3rd to the 6th decile it was predicted to be a non-excess return firm and if it was in the 7th or greater decile it was predicted as an excess return firm.

In period 2 the final market-adjusted disjunct for non-excess returns was based on an operating income to total assets ratio in the lowest 6 deciles and a change in earnings not in the top two deciles. The final disjunct for excess returns was book value to market value ratio between the 3rd and 9th deciles, change in earnings in the 5th or greater decile, and earnings to total asset ratio between the 4th and 9th deciles. The final disjuncts for size-adjusted returns were for non-excess firms with an operating income to total asset ratio in the bottom three deciles. The final disjunct for excess firms was an operating income to total assets ratio in the 3rd or greater decile, with a depreciation to property plant and equipment ratio in the 5th decile or less, and a percentage change in sales between the 2nd and 8th decile.

An interpretation of the trading rule for period 2 size-adjusted returns, for non-excess firms is as follows: A very low operating income to total assets ratio indicates that a firm is getting a poor return on its investment and therefore the shareholders will not earn excess returns on their investments. The trading rule for period 2 size-adjusted returns for excess firms is interpreted as follows: If the firm is getting an average return on its investments, it's sales are not changing drastically and the depreciation to property, plant and equipment ratio is not above average there is a good chance the shareholder will earn an excess return. Investors should be able to comprehend these rules. This is an example of the additional explanatory power possible with the genetic algorithm approach.

In an effort to determine if the fitness score is an appropriate indication of excess returns a comparison was made based on the fitness score and the associated return. If the fitness score was appropriate the higher fitness scores should have generated the greatest excess returns. The lowest fitness scores were for market-adjusted returns in period one. The total return for that model was .001667, the lowest value of any period. The next lowest fitness scores were for period one size-adjusted returns. That model had the next lowest return of .063173. The fitness scores for period two for both non-excess returns are higher than the fitness scores for excess returns in the same period. The short position, non-excess return firms, was the driving force behind the excess returns in this time period. Therefore, the ordering of the fitness scores corresponded to the ordering of the excess returns. This indicated that the fitness score was an appropriate measure.

5.4 Industry Specific Logistic Regressions

The next item tested was the use of industry specific logistic regressions to predict excess returns. A paired t-test, using the average excess returns for the industry-specific models, U_{ilr}, and the market-wide models, U_{lr}, for each year from 1983 to 1992 tests this hypothesis. The market-wide models were significantly better than the industry specific models. See table 11 for the results of the t-test. These results indicate that the null hypothesis can not be rejected. Using industry-specific models does not appear to increase the predictive accuracy of the models. One explanation of these results could be that the small number of firms in each model prevented an accurate mapping of the market returns. The formal hypothesis is shown below.

H_3: The returns from the industry-specific models will be less than or equal to the returns from the market-wide logistic regression models.

H_{3a}: The returns from the industry-specific models will be greater than the returns from the market-wide logistic regression models.

Table 11 - T-test for H_3

	Number of Firms in Short Portfolio	Short Return	Number of Firms in Long Portfolio	Long Return	Total Return	Return Type
Market-Wide Logistic Regression	5,096	-0.153411	5,096	0.022192	0.175603	Size Adjusted
	5,517	-0.129537	5,517	0.010907	0.140444	Market Adjusted
Industry-Specific Logistic Regression	5,741	-0.043808	5,741	0.022372	0.066180	Size Adjusted
	5,636	-0.018208	5,636	-0.006888	0.011320	Market Adjusted
One tailed T-test value for Size Adjusted Returns					0.010915	
One tailed T-test value for Market Adjusted Returns					0.003836	

Only those variables present in more than 60 percent of the firms were used in order to have an appropriate sample size for each industry as proxied by the two-digit SIC code. If the variables present in 60 to 70 percent of the firms were not significant then they were eliminated and the stepwise regression was rerun. Additionally, some variables which were constant for certain industries, primarily the LIFO and AQ dummy variables, were eliminated. The complete listings of the variables not included in each industry are in appendix tables A7 and A8. The complete listing of the industry-specific regression models developed after removing the variables listed above are in appendix tables A9 and A10.

Each regression model was used to generate a P_r score in each of the following five years for all firms that had each of the variables used in its particular industry-specific regression. The trading strategy was then applied on the pooled firms for each year. The complete results by year and position are in appendix tables A18 and A19.

The small population size of some industries may have reduced the impact of the industry-specific information. To determine if this explanation has some validity only those industries with a model probability of < .01 and more than 150 firm years were used. Regression models were developed and a trading strategy applied for each of these industries. A list of these industries is included in appendix table A11. This adjustment did not change the findings; the market-wide model still generated significantly greater returns. The complete results of this trading strategy are in appendix tables A22 and A23. However, the returns were still significantly less than the market-wide models' returns.

The study empirically investigated if neural networks and dual-layered genetic algorithms modeled the nonlinearity of markets better than logistic regression. The study also investigated if industry-specific models captured the environmental factors better than market-wide models. This study helps determine the most appropriate modeling method to predict excess returns considering market nonlinearity and environmental factors.

The results indicated that there was a significant difference in returns for each hypothesis. However the direction of the returns was opposite of the alternative hypotheses. Therefore the formal null hypotheses were not rejected. The implications of these results are discussed in the following chapter.

CHAPTER 6

CONCLUSIONS, LIMITATIONS, AND IMPLICATIONS FOR FUTURE RESEARCH

6.0 Introduction

The first section in this chapter reviews the improved results of the market-wide logistic regression models. The returns generated in this study are greater than the returns reported by Holthausen and Larcker and Ou and Penman. The rejection of the hypothesis that neural network models will produce greater returns than market-wide logistic regression models is then discussed. The next section reviews the rejection of the hypothesis that genetic algorithms would produce greater returns than logistic regression models. A discussion of the industry-specific models follows. Finally, implications for future research are discussed.

6.1 Conclusions and Limitations

This study confirms that the market-wide logistic regression is the most appropriate model to use for a trading strategy based on predicting excess or non-excess returns. See tables 10, 11 and 12 for the returns from each strategy. All of the alternative hypotheses were rejected at the .05 level as indicated in tables 10, 11 and 12

The results show that the initial classification of variables into four major groupings has merit. In both of the second period's final models using the market-wide logistic regression at least one variable was significant from each group. In the first period the size-adjusted return model had variables from each group except the other classification. The market-adjusted return model had variables from each group except the size classification. This is an unexpected finding since size was only significant when it was adjusted for in the returns. This indicates that the adjustment of returns based on size decile may compound the size anomaly.

Another item to review in the models is their predictive accuracy. The accuracy rates for both excess and non-excess returns are shown in appendix tables A24 to A27. The market-wide logistic model had the best predictive accuracy as well as the best trading strategy return. This suggests that the market-wide logistic model is the most appropriate model.

6.1.1 Prior Studies

One possible explanation for the difference in the returns between the current and prior studies is that during the 1973 to 1983 time period there may have been a smaller variance in returns between firms, since the time periods are not overlapping. The trading strategy is attempting to locate those firms that differentiate themselves from the normal. Another possible explanation is that by selecting certain industries the model incorporates more similar environmental factors. This lacks some support due to the results of the industry-specific models. However, using the two-digit DNUM may not be the most appropriate grouping of firms.

Another possible explanation is the different variables available. This study used a guided search that included a measure for labor force size and the firm's capital expenditure, which were in some of the final logistic regression models. These variables were both selected based on the importance noted by market analysts, and were not candidate variables in the prior studies. Both models in the second time period included the LIFO dummy variable incorporated from Lev and Thiagarajan's study.

Abarbanell and Bushee (1998) developed a trading strategy based on Lev and Thiagarajan's signals or variables. They found a potential for profit taking due to the market's under utilization of the information contained in the variables. The addition of these variables contributed to a more robust model. Each of the final models had at least one variable from Lev and Thiagarajan's variables, Ou and Penman's variables and Holthausen and Larcker's variables.

The size variables added to control for risk proxies, MV and BVMV, were significant in the second period only. The first trading period generated the greatest returns indicating that size did not drive the results.

The use of more extreme tails of the population may be a factor in the increase in performance from this study's logistic regression trading strategy. Holthausen and Larcker used 30 percent of the firms to invest in short and long. Ou and Penman used 40 percent of the firms for each investment position.

The modifications in the model development from Holthausen and Larcker appear to be useful predicting tools. The guided variable search appears to be more appropriate than a blanket type approach. An improvement in the predictive accuracy in the testing period was achieved by using the predictive accuracy as a guideline for model selection. The following sections review the hypothesis on an individual basis.

6.1.2 Neural Networks

The returns from using a trading strategy based on neural network models did not generate greater returns than the market-wide logistic regression models. This may indicate that neural networks do not adequately map complex market phenomena. The first null hypothesis was not rejected. A discussion of this failure to reject the null hypothesis follows.

This failure to reject the null hypothesis indicates that the noise in the training patterns does not allow for a proper mapping using the sigmoidal or tanh functions. It appears that the neural network methodology may not be appropriate for financial market return prediction. The noise in the financial markets may require a more advanced transfer function than the tanh and sigmoidal functions.

Ruggiero (1997a) argues that the root mean squared error function allows larger errors on rare large moves. This would then account for large losses on some trades limiting the effectiveness for backpropagation neural networks in developing profitable trading strategies. This is an additional argument for the use of the binary variables in developing the model. The use of a binary variable will help reduce the problem of rare large moves having a disproportionate influence on the model development.

As with any neural network modeling problem, one can never state that the correct model was found. Since the initial weight matrix is generated randomly and influences the success of the model, the most appropriate weight matrix may never be created. Different combinations of layers and nodes could develop a more accurate model. Using different methods of modifying the weight matrix or transformation functions could improve the neural network performance. Modifying the attributes used in this study could develop a more appropriate neural network model and therefore generate greater excess returns.

To properly train neural networks requires a "large and representative sample of relatively clean training data." (Kahn and Basu 1994) The market data may have too much noise to create a proper neural network model.

Another possible problem is the time frame used. This study developed neural networks on an annual basis. Hill, O'Connor andRemus (1996) compared the performance of neural networks to deseasonalized single exponential smoothing, Box-Jenkins, graphical (human judgment), naïve and deseasonalized Holt's exponential smoothing for various time series data. The neural network's performance was mixed in their study when the data were on an annual basis. However when quarterly series were used the neural network model performed significantly better than all other models. This may indicate that annual data contains too much noise and the current results may not hold if quarterly data is used.

Deco, Neuneier and Schurmann (1997) also indicate that neural networks do not work well with noisy financial time series. They predicted the daily differences of the German stock index, DAX, with a neuro-fuzzy approach that transforms rules given by experts into parameters that initialize a neural network. Using this modified approach could only produce mixed results. The profitability of a trading system based on the neural network only outperformed a naïve trading strategy for part of the test period.

6.1.3 Dual-Layered Genetic Algorithms

The returns from using a trading strategy based on genetic algorithm models did not generate greater returns than the logistic regression model. Therefore the second null hypothesis was not rejected. A discussion of this result follows. The dual layered genetic algorithm approach showed some promising results. In three of the four trading periods excess returns of more than .06 were achieved. Additionally, the disjuncts used allowed for an understanding of the trading strategy. A more appropriate fitness function could improve on the results achieved by allowing for the use of multiple disjuncts to select firms for the long and short positions.

The genetic algorithm model may not have found the best disjunct. An improved disjunct may be possible for the genetic-algorithm approach using an exhaustive search. This would ensure development of the best disjunct but at the expense of additional computational time. One way to limit the search would be to look at only four variable combinations or less since the best disjuncts found never contained more than four variables. This reduction in search space would save computational time but would no longer ensure that the best disjunct is developed. Another possible limitation on the search space would be to use all possible combinations of the variables found to be significant using logistic regression. This may help to add explanatory power to the logistic regression results.

Another possible problem is in the development of the fitness function. The current study looked for one disjunct to cover as many positive or negative examples as possible. If a fitness function were developed that looked only for disjuncts that covered positive examples it would limit the number of firms purchased long or sold short and may develop a better investment strategy. An example of this is in the period two size-adjusted returns model where the final disjunct was only part of the best KnowledgeSeeker rule. This study was looking at covering approximately half of all firms with two disjuncts. Another possibility is to use data in smaller intervals such as each five percent of firms instead of 10 percent. One could also permit non-continuous ranges of data in the disjuncts but this causes a loss of explanatory power. It would be difficult to justify a disjunct that included a rule such that if a value is in the 1st, 3rd or 7th decile then it will have excess returns.

Genetic algorithms should be researched further due to the additional explanatory capabilities they provide. Matthews (1996) describes an instance where Conrad Fellowman of SearchSpace developed a neural network model for bankruptcy prediction with a reported accuracy rate of 75 percent. They were unable to sell this information to the client due to the lack of justification of the model. SearchSpace developed a model, with a lower accuracy rate, using genetic algorithms that the client purchased due to their ability to comprehend the rules from the genetic algorithm.

6.1.4 Industry-Specific Logistic Regression

The returns from using a trading strategy based on industry-specific logistic regression models did not generate greater returns that the market-side logistic regression models. Therefore the third null hypothesis was not rejected.

A discussion of this failure to reject the null hypothesis follows. The industry-specific model returns for all industries were the lowest of all the original methodologies. An explanation for this is the small number of firms in some industries. When a reduced sample of industries was used, the returns for the market-adjusted model improved from .011320 to .072503. Another possibility is that Greig (1992) is correct in alleging that the ratios are factoring in industry effects. The market-wide regression model may be selecting the firms by using the ratios that best distinguish between industries to allow for purchasing long those industries that are strong and selling short those industries that are performing poorly.

Eliminating the industries with fewer than 150 firms on the Compustat PC Plus CDROM may have increased the ability of the P_r value to detect market sectors. The reduction in firms from small industries may have eliminated noise from those firms that may not have fit any general pattern. Cottle et. al. (1988) state that "analysts recognize that the stock market consists of numerous large market sectors or segments that-in terms of group price movements and total return results-are characterized by disparate performance." This study may have found a way of grouping firms into appropriate sectors for investment purposes.

The variables used in the models varied by industry. This supports the assumption that additional information can be gained by using industry-specific models. However, the P_r score's best function appears to be the selection of the appropriate industry for investment. The P_r score may allow for the selection of industries to invest in long or short based on standard industry ratios. By having the P_r score for each firm based on an industry-specific ratio, a separation is forced within each industry. Therefore, the individual stocks may show more volatility than the industry as a whole. When using the market-wide regression the P_r score can segregate firms by industry and thus obtain greater returns. The regression models for each industry and time period are in appendix tables A9 and A10.

6.2 Implications for Future Research

The primary driving force in all of the excess returns is the selling short position as is consistent with other market studies. Maybe part of the problem is in how excess returns are calculated. Looking at equation (1):

$$MAR_{im} = \sum_{t=1}^{m} (1 + R_{it}) - \sum_{t=1}^{m} (1 + R_{Mt})$$

a negative return is one that is less than the market return. If the market return is substantial then a large negative return may just be a failure of the firm's return to match the market return. If an individual actually held the short position the firm may not be decreasing in price and the amount of the return to the investor would not be as great as indicated. Industry-specific logistic regression models appear to be too sensitive to the training period due to the smaller number of firm years using a two-digit DNUM classification. An alternate approach may be to use one-digit DNUM classification.

The models were developed and used over a five-year time period without adjusting the model on a more frequent basis. The use of a moving time frame may improve the predictive ability of the models. An indicator of this potential problem would be a decreasing return during the five-year testing period. There does not appear to be any pattern in the reduction of the models' returns over time, but this is not formally addressed and may merit further review.

Another result with implications for future studies is that more size variables were included in the size-adjusted return models than in market-adjusted return models. This indicates that the size effect is not handled properly by using size-adjusted return. Further research needs to be undertaken to determine an appropriate method of accounting for size in market studies. This is a universal phenomenon that needs to be adequately addressed.

Neural networks should be developed with quarterly data to determine if Hill, O'Connor and Remus's (1996) findings can be supported. If this shows that quarterly data allows for the development of a more effective trading strategy, then future studies should make use of quarterly data if available.

In summary the results from all of the modeling techniques indicate that abnormal returns can be earned. The market-wide logistic regression models appear to develop the best trading strategies to generate excess returns. The changes in model development from prior studies to a guided search using classification accuracy as the primary model selection criteria generated greater returns than the prior studies. The noise in financial markets does not appear to be mapped appropriately by neural networks using the sigmoidal or tanh functions. The dual-layered genetic algorithm approach warrants further review due to the explanatory capability of the models.

Bibliography

Abarbanell, J. and B. Bushee, 1998, "Abnormal Returns to a Fundamental Analysis Strategy," The Accounting Review, Vol, 73, No. 1, January, pp. 19-45.

Abraham, A., and D. L. Ikenberry, 1994, "The Individual Investor and the Weekend Effect," Journal of Financial and Quantitative Analysis, Vol. 29, No. 2, pp. 263-277.

Agrawal, A. and K. Tandon, 1994, "Anomalies or Illusions? Evidence from Stock Markets in Eighteen Countries," Journal of International Money and Finance, Feb, v13 n1 pp. 83-107.

Ahmed, Anwer S., 1994, "Accounting Earnings and Future Economic Rents," Journal of Accounting and Economics, 17, pp. 377-400.

Ali, A. and P. Zarowin, 1992, "The Role of Earnings Levels in Annual Earnings-returns Studies," Journal of Accounting Research, Autumn, pp. 286-296.

Ali, Ashiq, 1994, "The Incremental Information Content of Earnings, Working Capital from Operations and Cash Flows," Journal of Accounting Research, Vol, 32, No. 1 Spring, pp. 61-74.

Avouyi-Dovi, S. and R. Caulet, 1995, "Using Artificial Neural Networks to Forecast Prices of Financial Assets," in Intelligent Hybrid Systems, ed., S. Goonatilake and S. Khebbal, New York, Wiley, pp. 360-377.

Bain, J.S., 1956, Barriers to New Competition, Cambridge, Mass.: Harvard University Press.

Ball, R., 1992, "The Earnings-price Anomaly," Journal of Accounting and Economics 15, pp. 319-345.

Ball, R. and S.P. Kothari, 1991, "Security Returns Around Earnings Announcements," The Accounting Review, Vol. 66, pp. 718-738

Bansal, A. R. J. Kauffman, and R.Weitz, 1993, "Comparing the Modeling Performance of Regression and Neural Networks as Data Quality Varies: A Business Value Approach," Journal of Management Information Systems, Vol. 10, No. 1, pp. 11-32.

Bauer, Richard, 1994, Genetic algorithms and Investment Strategies, New York: Wiley.

Bentz, Y., Refenes, A.N. and J.F. De Laulanie, 1995, "Modelling the Performance of Investment Strategies. Concepts, Tools and Examples," in Intelligent Hybrid Systems, ed., S. Goonatilake and S. Khebbal, New York, Wiley, pp. 241-258.

Bernard, V.L. and J.K. Thomas, 1990, "Evidence that Stock Prices do not Fully Reflect the Implications of Current Earnings for Future Earnings," Journal of Accounting and Economics, Vol. 13, p. 305-340.

Bhushan, Ravi, 1994, "An Informational Efficiency Perspective on the Post-earnings Announcement Drift," Journal of Accounting and Economics, 18, pp. 45-65.

Biddle, G. and G. Seow, 1991, "The Estimation and Determinants of Association Between Returns and Earnings: Evidence from Cross-Industry Comparisons," Journal of Accounting, Auditing and Finance, pp. 183-232.

Biggs, D., DeVille, B. and E. Suen, 1991, "A Method of Choosing Multiway Partitions for Classification and Decision Trees," Journal of Applied Statistics, Vol. 18, No. 1, pp. 49-62.

Bolland, P.J. and A.N. Refenes, 1994, "Analysis of the Relationship Between Volume, Open Interest and Future Prices," Proceedings of Neural Networks in the Capital Markets, Nov. 16-18.

Brock, W.A., D. Hsieh and B. LeBaron, 1991, Nonliner Dynamics, Chaos, and Instability Statistical Theory and Economic Evidence, MIT Press, Cambridge, MA.

Brorsen, B.W. and S. Yang, 1994, "Nonlinear Dynamics and the Distribution of Daily Stock Index Returns," The Journal of Financial Research, Vol. 17, No. 2 Summer, pp. 187-203.

Chan, L.K.C. and J. Lakonishok, 1992, "Are the Reports of Beta's Death Premature," Journal of Portfolio Management, 19, pp. 51 - 62.

Chan, L.K.C., Y. Hamao and J. Lakonishok, 1992, "Fundamentals and Stock Returns in Japan," Journal of Finance, 46, pp. 1739-1764.

Chiang, W.C., T.L. Urban and G.W. Baldridge, 1996, "A Neural Network Approach to Mutual Fund Net Asset Value Forecasting," Omega, Vol. 24, No. 2, pp. 205-216.

Clarke, R., 1989, "SICs as Delineators of Economic Markets," Journal of Business, January, pp. 17-31.

Collins, D. and S.P. Kothari, 1989, "An Analysis of Intertemporal and Cross-sectional Determinants of Earnings Response Coefficients," Journal of Accounting and Economics, 11, pp. 143-181.

Colin, Andrew, 1994 "Genetic Algorithms for Financial Modeling," in Trading on the Edge : Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets, ed. Guido J. Deboeck. New York: Wiley pp. 148 - 173.

Corhay, A. and Tourani Rad A., 1994, "Statistical Properties of Daily Returns: Evidence from European Stock Markets," Jounal of Business Finance & Accounting, Vol. 21, No. 2, pp. 271-282.

Cottle, S., R. Murray, and F. Block, 1988, Graham and Dodd's Security Analysis, McGraw-Hill, New York.

Davis, L., 1991, Handbook of Genetic Algorithms, Van Nostrand Reinhold, New York.

Deco, G., R. Neuneier and B. Schurmann, 1997, "Non-parametric Data Selection for Neural Learning in Non-stationary Time Series," Neural Networks, vol. 10, No. 3, pp. 401-407.

Donaldson, R.G., and M. Kamstra, 1996, "Forecast Combining with Neural Networks," Journal of Forecasting, Vol. 15, No. 1, pp. 49-52.

Downs, Thomas, 1991, "An Alternative Approach to Fundamental Analysis: The Asset Side of the Equation," The Journal of Portfolio Management, Winter, pp. 6-16.

Dutta, S., and S. Shekhar, 1988, "Bond Rating: a Non-conservative Application of Neural Networks," Proceedings of the Second International Conference on Neural Networks, San Diego, pp. 443-450.

Easton, Peter, 1985, "Accounting Earnings and Security Valuation: Empirical Evidence of the Fundamental Links," Journal of Accounting Research, 23 Supplement, pp. 54-77.

Easton, P. and T. Harris, 1991, "Earnings as an Explanatory Variable for Returns," Journal of Accounting Research, Spring, pp. 19-36.

Eberhart, R.C., 1992, "The Role of Genetic Algorithms in Neural Network Query-Based Learning and Explanation Facilities," Proc. Intl. Joint Conf. Neural Networks, pp. 169-182.

Edmonds, A., D. Burkhardt, and O. Adjei, 1995, "Genetic Programming of Fuzzy Logic Production Rules with Application to Financial Trading," in Intelligent Hybrid Systems, ed., S. Goonatilake and S. Khebbal, New York, Wiley, pp. 179-188.

Ely, Kirsten M., 1991, "Interindustry Differences in the Relation Between Compensation and Firm Performance Variables," Journal of Accounting Research, Vol. 29, No. 1 Spring, pp. 37-58.

Fama, E.F. and K.R. French, 1992, "The Cross-section of Expected Stock Returns," Journal of Finance, 47, pp. 427-465.

Fogler, H. R.,1995, "Investment Analysis and New Quantitative Tools," The Journal of Portfolio Management, Summer , pp. 39-48.

Finger, Catherine A., 1994, "The Ability of Earnings to Predict Future Earnings and Cash Flow," Journal of Accounting Research, Vol. 32, No. 2 Autumn, pp. 210-223.

Freeman, R. N. and S. Tse, 1992, "A Nonlinear Model of Security Price Responses to Unexpected Earnings," Journal of Accounting Research, Vol. 30, No. 2 Autumn, pp. 185-209.

French, K.R., 1980, "Stock Returns and the Weekend Effect," Journal of Financial Economics, Vol. 8, pp. 55-69.

Gangopadhyay, P., 1994, "Risk-Return Seasonality and Macroeconomic Variables," The Journal of Financial Research, Vol. 17, No. 3. pp. 347-361.

Gencay, 1995, "Forecast Performance of Moving Average Rules with Stock Returns," in Intelligent Hybrid Systems, ed., S. Goonatilake and S. Khebbal, New York, Wiley, pp. 215-226.

Giles, C.L. and C.W. Omlin, 1994, "Pruning Recurrent Neural Networks for Improved Generalization Performance," IEEE Transactions on Neural Networks, Vol. 5, No. 5, September, pp. 848-851.

Greig, Anthony, 1992, "Fundamental Analysis and Subsequent Stock Returns," Journal of Accounting and Economics, 15, pp. 413-442.

Haefke, Christian and Helmenstein, Christian 1994. "Stock Price Forecasting for Austrian Initial Public Offerings Using Artificial Neural Networks," Proceedings of Neural Networks in the Capital Markets, Nov. 16-18.

Hall, James, 1994, "Adaptive Selection of U.S. Stocks with Neural Nets," in Trading on the Edge : Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets, ed. Guido J. Deboeck. New York: Wiley pp. 45 - 65.

Hand, J., 1990, "A Test of the Extended Functional Fixation Hypothesis," The Accounting Review, Vol. 65, pp. 740-763.

Hand, J.R.M., 1991, "Extended Functional Fixation and Security Returns Around Earnings Announcements: a Reply to Ball and Kothari," Accounting Review, Vol.66, No. 4, pp. 739-747.

Hildebrand, D. K. And L. Ott, 1987, Statistical Thinking for Mangers, PWS, Boston, MA.

Hill, T., M. O'Connor, and W. Remus, 1996, "Neural Network Models for Time Series Forecasts," Management Science, July, Vol. 42, No. 7, pp. 1082-1093.

Hintze, J. L. 1995, NCSS 6.0 , NCSS, Kaysville, UT

Holland, J.H., 1992, "Genetic Algorithms," Scientific American, July, pp. 66-72.

Holthausen, R.W. and D. Larcker, 1992, "The Prediction of Stock Returns using Financial Statement Information," Journal of Accounting and Economics, 15, pp. 373-411.

Hsieh, D., 1991, "Chaos and Nonlinear Dynamics: Applications to Financial Markets," Journal of Finance, 46, pp. 1839-1878.

Janikow, C.Z., 1993, "A Knowledge-Intensive Genetic Algorithm for Supervised Learning," Machine Learning, 13, pp. 189-228.

John, G.H., P. Miller, and R. Kerber, 1995, "Stock Selection using RECON," in Intelligent Hybrid Systems, ed., S. Goonatilake and S. Khebbal, New York, Wiley, pp. 303-316.

Kahn, R.N. and A.K. Basu, 1994, "Neural Networks in Finance: An Information Analysis," Proceedings of Neural Networks in the Capital Markets, Nov. 16-18.

Kass, G.V., 1980, "An Exploratory Technique for Investigating Large Quantities of Categorical Data," Applied Statistics, Vol. 29, No. 2, pp. 119-127.

Keim, D.B., 1983, "Size-related Anomalies and Stock Return Seasonality Further Empirical Evidence," Journal of Financial Economics, Vol. 12, pp. 13-32.

Keim, D.B. and R.F. Stambaugh, 1984, "A Further Investigation of the Weekend Effect in Stock Returns," The Journal of Finance, Vol. 39, No 3, pp. 819-835.

Kiernan, V., 1994, "Growing Money from Algorithms," New Scientist, Vol. 144, No. 1954, pp. 25-28.

Kim, C. and J. Park, 1994, "Holiday Effects and Stock Returns: Further Evidence," Journal of Financial and Quantitative Analysis, Vol. 29, No. 1, pp. 145-157.

Kramer, Charles, 1994, "Macroeconomic Seasonality and the January Effect," The Journal of Finance, Vol. 49. No. 5, pp. 1883-1891.

Kuan, C., and White, H., 1994, "Artificial Neural Networks: An Econometric Perspective," Econometric Reviews, Vol. 13, No. 1, pp. 1-91.

Lam, K.P., K.C. Chiu and W.K. Chan, 1995, "An Embedded Fuzzy Knowledge Base for Technical Analysis of Stocks," in Intelligent Hybrid Systems, ed., S. Goonatilake and S. Khebbal, New York, Wiley, pp. 317-330.

LeFebvre, W., 1995, NeuroSolutions Reference Manual, NeuroDimensions Inc., Gainesville, FL.

Lev, Baruch, and S. Ramu Thiagarajan, 1993, "Fundamental Information Analysis," Journal of Accounting Research, Vol. 31, No. 2, Autumn, pp. 190-215.

Levitt, Marc, 1994. "Improved Exploitation of Market Inefficiencies by Nonlinear Forecasting Techniques," Proceedings of Neural Networks in the Capital Markets, Nov. 16-18.

Lustgarten, S. and S. Thomadakis, 1980, "Valuation Response to New Information: A Test of Resource Mobility and Market Structure," Journal of Political Economy, October, pp. 977-993.

Malliaris, M.E., 1994, "Modeling the Behavior of the S&P 500 Index: A Neural Network Approach," IEEE Transactions in Neural Networks, pp. 86-89.

Matthews, Robert, 1996, "Panning for Data Gold," New Scientist, May 25, Vol. 150, No. 2031, pp. 30-34.

Morgan, J.A. and Sonquest, J.N., 1963, "Problems in the Analysis of Survey Data: and a Proposal," Journal of American Statistical Association, Vol. 58, pp. 415-34.

Myers, R.H., 1990, Classical and Modern Regression with Applications, PWS-Kent, Boston, MA.

Ou, J. and S. Penman, 1989, "Financial Statement Analysis and the Prediction of Stock Returns," Journal of Accounting and Economics, 11, pp. 195-329.

Palepu, K., V. Bernard, and P. Healy, 1996, Business Analysis & Valuation, Cincinnati, OH.

Porter, M., 1980, Competitive Strategy, The Free Press, New York, NY

Reed, R., 1993, "Pruning Algorithms - A Survey," IEEE Transactions on Neural Networks, Vol. 4, No. 5 September, pp. 740-747.

Refenes, Apostolos N., Zapranis, Achileas and Francis, Gavin, 1994, "Stock Performance Modeling Using Neural Networks: A Comparative Study with Regression Models," Neural Networks, Vol. 7, No. 2, pp. 375-388.

Robles, J.V. and C.D. Naylor, 1995, "Applying Neural Networks in Copper Trading: Weekly Transaction ANN Model with a Linear Filter," in Intelligent Hybrid Systems, ed., S. Goonatilake and S. Khebbal, New York, Wiley, pp. 354-359.

Ruggiero, M.A., 1994, "How to Build an Artificial Trader," Futures, Vol. 23 No. 10 pp. 56-59.

Ruggiero, M.A., 1995, "Evolution of a Trader," Futures, Vol. 24, No. 7, pp. 44-47.

Ruggiero, M.A., 1997a, "Breeding a Super Trader," Futures, Vol. 26, No. 1, pp. 52-55.

Ruggiero, M.A., 1997b, "Evolution of a Super Trader," Futures, Vol. 26, No. 2, pp. 44-48.

Scheinkman, J.A. and B. LeBaron, 1989, "Nonlinear dynamics and stock returns," Journal of Business, 62, pp. 311-38.

Schwert, G.W., 1989, "Size and Stock Returns and other Empirical Regularities," Journal of Financial Economics, pp. 3-11.

Sen, Tarun K., R. Oliver, and N. Sen, 1995, "Predicting Corporate Takeover Targets: a Comparative Study of Neural Networks and Logistic Regression," Neural Networks in the Capital Markets, ed. A. N. Refenes, Wiley, pp. 325-340.

Shevlin, Terry and D. Shores, 1993, "Firm Size, Security Returns and Unexpected Earnings: The Anomalous Signed-Size Effect," Contemporary Accounting Research, Vol. 10, No. 1 Fall, pp. 1-30.

Shirreff, D., 1994, "Can Technology Beat the Markets?," Euromoney, Nov., pp. 30-35.

Sikora, R. and Shaw, M., 1994, "A Double-layered Learning Approach to Acquiring Rules for Classification: Integrating Genetic Algorithms with Similarity-Based Learning", ORSA Journal on Computing, Vol. 6, Spring, pp. 174-187.

Steiner, M. and H. Wittkemper "Asset Allocation with a Neural Network Implementation of the Coherent Market Hypothesis," Proceedings of Neural Networks in the Capital Markets, Nov. 16-18.

Wang, Yong, 1993, "Near-rational Behavior and Financial Market Fluctuations," Economic Journal, Vol. 103, No. 421, pp. 1462-1479.

White, H., 1988, "Economic Prediction using Neural Networks: The Case of IBM Daily Stock Returns," Proceedings of Second IEEE International Conference on Neural Nets, San Diego, pp. 451-458.

Yao, J., and H-L. Poh, 1995, "Equity Forecasting: A Case Study on the KLSE Index," in Intelligent Hybrid Systems, ed., S. Goonatilake and S. Khebbal, New York, Wiley, pp. 341-353.

Ye, Z. and Gu, L., 1994, "A Fuzzy System for Trading the Shanghai Stock Market," in Trading on the Edge : Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets, ed. Guido J. Deboeck. New York, Wiley, pp. 207-214.

Yoda, Mario, 1994, "Predicting the Tokyo Stock Market," in Trading on the Edge : Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets, ed. Guido J. Deboeck. New York, Wiley pg. 66-79.

Zapranis, A.D. and A.N. Refenes, 1994, "Neural Networks in Tactical Asset Allocation: Towards a Methodology for Hypothesis Testing and Confidence Intervals," Proceedings of Neural Networks in the Capital Markets, Nov. 16-18.