JOURNAL OF EMERGING TECHNOLOGIES IN ACCOUNTING
Vol. 5
2008
pp. 129–142

Restatements Due to Improper Revenue Recognition: A Neural Networks Perspective

Srinivasan Ragothaman
The University of South Dakota

Angeline Lavin
The University of South Dakota

ABSTRACT: The Securities and Exchange Commission (SEC) issued Staff Accounting Bulletin No. 101 (SEC 1999) in an attempt to curb improper revenue recognition practices. Nonetheless, revenue restatements and the subsequent earnings restatements have continued unabated. Our goal is to contribute to the emerging technologies literature by applying the neural networks methodology to the study of revenue restatements. We also compare the results of the neural network classification with classifications obtained from multiple discriminant analysis (MDA) and logistic regression (Logit) models. Six financial and governance variables were used to train the neural network on a sample of 180 firms, and the model was validated using a holdout sample of 51 additional firms. The results show that the neural network model has superior predictive power for predicting revenue restatement firms when compared to the MDA and Logit models. However, the Logit and MDA models predict nonrevenue restatement firms better. Moreover, when misclassification costs are included, the neural network (NN) model performs the best with the lowest relative misclassification costs.

Keywords: neural networks; revenue restatements; discriminant analysis; logistic regression; misclassification costs.

INTRODUCTION

Publicly traded firms operate under intense pressure to "make their numbers" each quarter. Firms that do not meet expectations are often severely punished by the market. This potential pressure creates an incentive for earnings management. Unfortunately, improper revenue recognition sometimes results when firms are pressured to meet investor expectations. Revenue recognition shenanigans have taken place for many years. The 1987 COSO study (National Commission on Fraudulent Financial Reporting [NCFFR] 1987) reported that in 47 percent of the fraud cases reviewed from 1981 to 1986, improper revenue recognition was an issue. The 1999 COSO study (NCFFR 1999) suggested that

We thank two anonymous reviewers, Miklos Vasarhelyi (editor), an associate editor, and the participants at the 2005 AI/ET National Workshop of the American Accounting Association for many useful comments. The usual disclaimer applies. This research has been partially supported by a grant to the first author from The University of South Dakota Research Catalyst Fund.

Corresponding author: Srini Ragothaman
Email: Srini.Ragothaman@usd.edu

improper revenue recognition persisted as a tool of financial misstatement. It was a factor in 50 percent of the fraud cases from 1987 to 1997. Corporate earnings restatements cost investors $100 billion from 1997 to 2002 (General Accounting Office [GAO] 2002). Of the 689 restatements examined by this GAO report, improper revenue recognition issues accounted for the largest number (39 percent) of earnings restatements. Furthermore, restatements for improper revenue recognition result in larger drops in market capitalization than any other type of restatement (GAO 2002).

The large volume of revenue restatements forced the Securities and Exchange Commission (SEC) to issue Staff Accounting Bulletin (SAB) No. 101 (SEC 1999), to curb improper revenue recognition practices. Although the adoption of SAB 101 was somewhat controversial, anecdotal and empirical evidence suggest that improper revenue recognition is an important issue that deserves attention (Altamuro et al. 2005). The matching principle of revenue recognition requires that revenue be recognized at the same time as costs and adjustments to sales are booked and when the transfer of title occurs (Guerra 2003). However, this principle is sometimes ignored when financial statements are prepared.

In response to the revenue manipulations, Congress passed the Sarbanes-Oxley Act (U.S. House of Representatives 2002), the SEC has increased their oversight and enforcement, law enforcement became involved with pursuing wrongdoers, and the American Institute of Certified Public Accountants (AICPA) issued the Statement on Auditing Standards No. 99 (SAS 99) in 2002. SAS 99, *Consideration of Fraud in Financial Statements*, has significantly increased auditors' responsibilities to provide reasonable assurance that financial statement amounts are free from material misstatements that emanate from fraud. Table 1 provides a brief description of some of the most common revenue manipulation schemes and is based on discussions in Stallworth and Digregorio (2004), GAO (2002), and Guerra (2004).

Researchers have been studying the characteristics of firms that restate earnings for many years using a variety of statistical techniques including regression analysis, correlation analysis, and discriminant analysis (Wu 2002; Palmrose et al. 2004; Byrne 2002; Palmrose and Scholz 2004; Owers et al. 2002). However, these statistical models use objective information only, ignoring relevant subjective information. Furthermore, statistical models assume multivariate normality and homoscedastic variances despite the fact that these assumptions are routinely violated with real world data (Etheridge et al. 2000). In light of these constraints with statistical models, computer-based decision support systems and neural networks (NN) have become increasingly popular for decision making in the fields of finance, accounting, and marketing. Neural networks have long been used to study bankruptcy, including the factors that may suggest impending bankruptcy and predict financial distress. There is evidence to suggest that under certain circumstances, neural networks perform better than standard statistical models and can be used as valuable decision aids (Shawver 2005; Ragothaman 2003; Ragothaman 2001; O'Leary 1998; Fanning et al. 1995). Thus, this research will focus on the development of a prototype neural network for identifying firms that are likely to restate revenues. This paper is the first step in using neural networks to study revenue recognition improprieties.

A key purpose of this research is to examine the characteristics common to firms that restate revenues and to use this information to develop a neural network model that will predict revenue restatements by corporations. Such a model should be of interest to a variety of constituents including auditors, financial analysts, investors, and government regulators. In particular, a neural network model to improve auditors' understanding of ''revenue restatement firms'' is an important contribution to the field of auditing. A model that could predict revenue improprieties would also be of value to investment analysts and portfolio

TABLE 1
Common Revenue Manipulation Schemes[a]

1.	Channel stuffing	Offering of incentives to wholesalers and retailers to take delivery of more inventory than needed for their immediate sales.
2.	Revenue recognized when there is no transfer of property	Recording revenue in the current quarter but shipping after the current quarter ends.
3.	Deferred accounting for discounts, rebates, etc.	Accounting for sales in the current period and deferring accounting for discounts, rebates, etc., to a future period.
4.	Sales without substance	Funding the buyer so that collection is assured.
5.	Side agreements	These side agreements contain cancellation provisions, right of return, product exchanges, refunds, etc., that are not accounted for.
6.	Round-trip transactions	Two Internet firms provide advertising for each other, no cash is exchanged, and fair value is not determinable.
7.	Fictitious sales recorded	Creating false documents to record fictitious sales.
8.	Bill and hold schemes	The customer agrees to buy and is billed, but the seller retains possession of inventory until the buyer requests delivery.
9.	Reporting revenues at gross	Recording revenues at gross rather than net. For example, travel agencies report gross revenues rather than net commissions.
10.	Extended payment terms	Payment terms granted far exceed the firm's normal terms. Thus, is collection probable?

[a] Sources: Stallworth and Digregorio (2004), GAO (2002), and Guerra (2004).

managers who make investment decisions based on earnings forecasts. A prediction model may be especially valuable in the current environment of more limited company disclosures due to the requirements of Regulation FD.

A brief review of the prior research in the areas of earnings and revenue restatements is presented in the next section. The third section describes the sample selection procedures and the data used in this study. The process of training the neural network is described in the fourth section. In the fifth section, the results from all three models as well as the holdout sample results are discussed. In the sixth section, a discussion of misclassification costs is offered. Last, a brief summary and limitations are presented.

PRIOR RESEARCH

The cases of MicroStrategy, Cendant, and Sunbeam illustrate the substantial market reaction that a few revenue restatements can cause. In the seven-day period around the announcement of their respective restatements, the three aforementioned firms together lost more than $23 billion in market value (Turner et al. 2001). When MicroStrategy announced a revenue restatement of 25 percent in March 2000, the stock price declined by $140 or 62 percent. A full description of what transpired at MicroStrategy is available in AAER 1350 (SEC 2000).

Callen et al. (2002) find that restating firms tend to have deteriorating financial performance, both in terms of income and cash flow, at the time of the restatement. Furthermore, restatement firms tend to have poor performance relative to their industry peers. This evidence is consistent with the use of aggressive accounting practices to manage earnings in an attempt to improve financial statements and hide problems. Abdolmohammadi and Read (2004) report that certain audit opinions are related to future financial restatements. According to Choi and Jeter (1992), audit qualifications indicate that uncertainties associated with future cash flows are increased, and consequently, the future market value of the firm can be adversely impacted.

There are several ways to approach the identification of firms that are most likely to experience revenue (and earnings) restatements. Some authors have approached the research question by studying the *structural aspects* of firms that have restated earnings while other researchers have focused on the *financial symptoms* that help diagnose the potential for restatements. Factors that separate firms with good corporate governance practices from firms with poor corporate governance practices may also be useful for predicting earnings restatements. Authors who have investigated the structural aspects include Dechow et al. (1996), who found that the likelihood of earnings manipulation is systematically related to weaknesses in management oversight such as the lack of an audit committee. More recently, Xie et al. (2003) show that composition of the audit committee is related to the likelihood that a firm will engage in earnings management. They find that more financially sophisticated board and audit committee members are associated with firms that have smaller current accruals, and they conclude that financially sophisticated board and audit committee members may play an important role in curtailing earnings management. Firm size may also be a predictor of the propensity of a firm to manage earnings. Kinney and McDaniel (1989) and DeFond and Jiambalvo (1991) found that firms making restatements are generally smaller firms. Lee and Choi (2002) argue that firm size is a proxy for information asymmetry. Managers of large companies are not able to retain private information as their counterparts in small companies can do with some success. They also report that small companies manage earnings more frequently than large firms to avoid losses.

The identification of financial symptoms associated with revenue and earnings restatements is really an attempt to study the earnings quality of firms. Schipper (1989) was one of the first to ‘‘frame’’ the inquiry into the area of earnings management. Richardson et al. (2002) undertook a comprehensive analysis of restatement firms. They found that restatement firms have very large accruals in the year of the alleged manipulation, higher market multiples (price-to-earnings as well as price-to-book ratios), and debt covenants (proxied by leverage and current ratio) that may provide motivation for aggressive accounting practices. Kinney and McDaniel (1989) found that firms making restatements are less profitable and tend to have higher leverage and declining earnings, which would suggest that these firms provide less return to shareholders. Wu (2002) found that 40 percent of restatements are due to either premature or false revenue recognition.

DATA AND SAMPLE SELECTION

This study is a first step toward studying the usefulness of neural networks in predicting revenue restatements. It focuses exclusively on earnings restatements caused by improper revenue recognition practices identified in the GAO (2003) study. We exclude restatements that occurred due to operational changes such as stock splits, currency-related issues, litigation settlements, arithmetic errors, and changes in accounting principles. We also exclude all restatements due to acquisitions and mergers, in-process R&D, related-party transactions, reclassifications, restructurings, mistakes in derivatives accounting, and misstatements in expenses. Of the 689 restatements analyzed by the GAO (2003), 39 percent were due to improper revenue recognition. Of these 268 revenue restatement firms, we had complete Compustat data and audit committee information for 110 revenue restatement (RR) firms.

This sample of 110 revenue restatement firms covers the period 1998 through 2002 (GAO 2003). Each of the restatement firms had one restatement during the sample period. Six financial and governance variables, which have theoretical support in published literature as described previously, were used in this study. Five of these six variables, which include risk and market return measures as well as financial ratios, were collected from the Compustat database. Specifically, the five independent variables (all continuous) collected from Compustat are the following: (1) *Current Ratio* (current assets divided by current liabilities—a liquidity measure); (2) *Total Return for One Year* (rate of return reflecting stock price appreciation plus reinvestment of dividends); (3) *Beta* (beta is a measure of the sensitivity of a company's stock price to the overall fluctuation in the S&P 500 index); (4) *Natural Logarithm of Total Assets* (a size measure); and (5) *Audit Opinion* (a corporate governance measure). The sixth independent variable is an indicator variable representing the financial expertise on the audit committee. The Audit Committee variable is coded 1 if a firm's audit committee contains at least one financial expert and is 0 otherwise. This is also a corporate governance measure. The audit committee data came from proxy statements. The data for the six variables were collected for one year prior to the announcement year.

The variables used in this study were chosen based on the findings of many of the studies discussed in the literature review. The current ratio captures the impact of debt covenants on management's decision to pursue aggressive accounting tactics. The log of total assets is a proxy for firm size. The audit committee variable is included because financially sophisticated audit committee members may play an important role in accounting policy choices. The total return variable is included because restatement firms tend to be less profitable and experience declining earnings, which generally leads to stock price declines. Beta, a measure of a stock's volatility relative to the market, is included to capture the likelihood that restatement firms experience more stock price volatility due to negative announcements. Continuous variables were used "as they are" while performing neural network analysis, Logit regression, and discriminant analysis. They were not converted into dichotomous variables.

A control sample of an additional 121 firms that did not experience revenue restatements (NRR firms) between 1998 and 2002 was randomly selected. Financial ratios and operating measures for the control sample were also obtained from the Compustat database and the proxy statements.

Descriptive statistics for the included variables are presented in Table 2. The mean of sales revenue for RR firms is 2.84 billion and 3.09 billion for NRR firms. The mean of total assets for RR firms is 3.23 billion and 2.92 billion for NRR firms. RR firms tend to be as large as NRR firms. In fact, the chi-squared statistic (1.449) in the Logit model for the size variable (*LNTA*) is not statistically significant. There is a higher concentration of financial experts on the audit committees of RR firms than NRR firms. RR firms tend to receive more qualified and modified audit opinions than NRR firms.

Industry classifications based on the two-digit SIC codes were examined. A variety of industries are represented in the sample, including pharmaceutical products and chemicals, retail and wholesale, business services, industrial machinery and equipment, electronic equipment, instruments, food, oil and gas, and a few others.

METHODOLOGY AND RESULTS

Neural Networks

Neural networks are capable of recognizing trends and patterns, learning from data, and then making predictions (Etheridge and Sriram 1997). Neural networks consist of nodes

TABLE 2
Descriptive Statistics

	RR Firms		NRR Firms	
	Mean	**Std Dev**	**Mean**	**Std Dev**
TRT1Y	24.68	92.06	−2.23	60.73
BETA	1.49	0.98	0.95	0.88
CR	3.48	7.29	2.29	1.36
LNTA	6.12	1.89	5.67	2.28
AC	0.73	0.45	0.46	0.50
AO	1.92	1.39	1.32	0.93
SALE	2.84	7.04	3.09	8.87
TA	3.23	8.68	2.92	7.54

Variable definitions:
TRT1Y = total return for one year in percentage terms;
BETA = market beta;
CR = current assets to current liabilities ratio;
LNTA = natural logarithm of total assets;
AC = 1 if the audit committee contains a financial expert, 0 otherwise;
AO = auditor's opinion in the audit report;
SALE = sales revenue in billions; and
TA = total assets in billions.

in multiple layers; these nodes are sometimes referred to as processing elements. The weight associated with each node denotes the degree of influence each node has on other nodes. Neural networks can be subjected to either unsupervised or supervised learning. In unsupervised learning, the network is presented with a set of input variables, but no output variable is provided to the network. The NN examines the input variables and formulates an algorithm to classify the data, which is similar to a multivariate technique that explores the underlying data structure (Boritz et al. 1995). Supervised learning occurs when the NN is presented not only with the input variables but also the correct output.

This paper describes the development, training, and testing of a neural network model using the commercially available neural network software, "Brainmaker" (Lawrence 1994; Hardgrave et al. 1994; Wilson and Sharda 1994; Ragothaman 2001; Barniv et al. 1997). In this study the Brainmaker software is used to create a standard back propagation network. Lawrence (1994) suggests that the back propagation network works very well with the sigmoid function. Error minimization is achieved by constantly modifying the input connection weights (Boritz et al. 1995). Back propagation learning algorithms have been used by Wilson and Sharda (1994) and other researchers for two-group classification models.

Fadlalla and Lin (2001) indicated that out of 40 finance and accounting NN studies reviewed, 29 reported using only one hidden layer. This study also uses a network with one hidden layer. There is little consensus among researchers regarding the number of nodes that a hidden layer should contain. Lawrence (1994) suggests that the hidden layer should contain one-half the nodes in the input layer. Wilson (1992) argues that the hidden layer can contain twice the number of input nodes plus one. Following Wilson (1992), 13 nodes were used for the hidden layer, and the input layer had six nodes for the six independent variables. Learning rate is basically a factor to scale all corrections while learning and helps the neural network to converge at a faster pace (Lawrence 1994). This study started with a learning rate of 3.6, and the rate was gradually reduced as the learning progressed to a final learning rate of 3 when the network converged. There is no accepted theory to guide the selection of training tolerance level.

Training tolerance is the acceptable error value, below which weights need not be updated (Barniv et al. 1997). If a tight training tolerance (e.g., 0.01, 0.02, 0.05, etc.) is specified, the result may be slow convergence and overtraining of the network. If the network is overtrained, it may not perform well on a holdout sample. If a large training tolerance (e.g., greater than 0.4) is specified, the network may not be sufficiently trained and may converge faster. A weakly trained NN also may not perform well on a holdout sample. Lawrence and Frederickson (1993) recommend starting with a large training tolerance (0.4) because it helps the network gain a general understanding of the overall problem. We started with a training tolerance of 0.35 and slowly reduced it to 0.33 and then 0.32, after some training. A final training tolerance of 0.25 was used. The "shuffle" utility available in Brainmaker was used to randomize the training set.

Results and Discussion

Table 3 shows the classification results of the six-variable neural network, MDA, and Logit models. Panel A gives the training sample results indicating that the six-variable neural network model classifies 83.5 percent of the RR and 50.5 percent of the NRR cases correctly.

Objectivity is desirable when validating an expert system or an AI model such as a neural network model (O'Leary 1987). One way to objectively validate a neural network model is to use different sets for development and testing. O'Leary (1987) also recommends validating AI models (machine learning algorithms, neural networks, and others) against other statistical models. Both validation methods are used to assess this neural network model's ability to classify firms into RR and NRR groups. Panel B shows that when the neural network model is used to analyze the holdout sample, 84 percent of the RR and 34.6 percent of the NRR firms are grouped correctly.

A multiple discriminant analysis (MDA) model and a Logit model are employed as content validation tools to evaluate the performance of the neural network model. The discriminant analysis was performed using SPSS, and the results are described in Table 3. The purpose of MDA is to determine the linear combination of the explanatory variables that best discriminates between groups that are partitioned. MDA is often applied to problems where the dependent variable is dichotomous, as in the current case. MDA classifies entries into mutually exclusive groups by maximizing the inter-group to intra-group variance-covariance from a set of predictor variables. Panel A of Table 3 gives the training sample results indicating that the six-variable discriminant model classifies 64.7 percent of the RR and 75.8 percent of the NRR cases correctly. Panel B shows that when the discriminant model is used to analyze the holdout sample, 60 percent of the RR and 76.9 percent of the NRR firms are grouped correctly. The holdout sample results indicate that while the neural network is significantly more accurate ($Z = 1.905$) in predicting RR firms, the MDA model is superior ($Z = 4.41$) in identifying NRR firms. The normally distributed Z-test for the equality of proportions was used (O'Leary 1998).

Much of the prior research in the area of financial NN provides a comparison of neural network results with a traditional linear technique such as MDA (Brown and Coakley 2000). Additional comparisons are made in this paper against a nonlinear technique. Logit analysis is designed to overcome the problems with applying Ordinary Least Squares (OLS) to a dichotomous dependent variable as in the current case. These problems result from the data violating the assumptions of OLS and invalidating the results of statistical tests and inferences. Furthermore, nonlinear Logit analysis relies on a less restrictive set of distribution

TABLE 3
Training and Holdout Sample Results

	Classification Matrix		Predicted Group TOTAL	
Actual Group	**% Correct**	**RR Firms**	**NRR Firms**	**TOTAL**
Panel A: Training Sample Results				
Neural Network				
RR Firms	83.5	71	14	85
NRR Firms	50.5	47	48	95
Type I Error: 16.5%* Type II Error: 49.5%				
MDA				
RR Firms	64.7	55	30	85
NRR Firms	75.8	23	72	95
Type I Error: 35.3%* Type II Error: 24.2%				
Logit				
RR Firms	58.8	50	35	85
NRR Firms	76.8	22	73	95
Type I Error: 41.2%* Type II Error: 23.2%				
Panel B: Holdout Sample Results				
Neural Network				
RR Firms	84.0	21	4	25
NRR Firms	34.6	17	9	26
Type I Error: 16.0%* Type II Error: 65.4%				
MDA				
RR Firms	60.0	15	10	25
NRR Firms	76.9	6	20	26
Type I Error: 40.0%* Type II Error: 23.1%				
Logit				
RR Firms	56.0	14	11	25
NRR Firms	73.1	7	19	26
Type I Error: 44.0%* Type II Error: 26.9%				

Type I (II) error is defined as the percentage of firms that were classified as NRR (RR) firms, while they were actually RR (NRR) firms.

assumptions than MDA. Specifically, linear discriminant analysis yields optimal solutions only when independent variables are drawn from a multivariate normal distribution and within-group variance-covariance matrices are identical. However, these assumptions are often violated in practice.

Similar to MDA, the fitted (predicted) values from the Logit model can be used to classify entries into mutually exclusive groups using a set of predictor variables. The Logit analysis was also performed using the SPSS software, and Table 3 shows the classification matrix obtained from the six-variable Logit (logistic regression) model. Panel A, which gives the training sample results, indicates that the Logit model classifies 58.8 percent of the RR and 76.8 percent of the NRR cases correctly. Panel B shows that when the Logit model is used to analyze the holdout sample, 56 percent of the RR and 73.1 percent of the NRR firms are grouped correctly. The holdout sample results indicate that while the neural network is significantly more accurate ($Z = 2.171$) in predicting RR firms, the Logit model is superior ($Z = 3.97$) in identifying NRR firms. Again, the normally distributed Z-test for the equality of proportions was used (O'Leary 1998).

Robustness tests were performed by including common financial measures, profit margin (*PM*), return on assets (*ROA*), and return on equity (*ROE*), as predictors instead of and in addition to the one-year market return variable in the MDA and Logit models. *ROA* and *PM* alone were each substituted for the market return variable, *ROA* and *ROE* together were substituted for the market return variable, and *ROE*, *ROA,* and *PM* were added to the models along with the market return variable. In all cases, the models using the one-year market return variable alone (as reported in Table 3) had a better overall accuracy rate than any of the other combinations or single variables. In addition, the combined Type I and Type II errors for the two statistical models in all of these robustness tests are slightly higher than what is reported in the paper using the one-year market return variable alone. The robustness test analysis indicates that the results reported in the paper using the one-year market return variable are more conservative.

Misclassification Costs

The comparison of the results of all three models shown in Table 3 suggests that the statistical models (MDA and Logit) are more accurate in predicting the NRR cases, while the nonlinear neural network model is more accurate than the statistical models in predicting the RR cases. Within the context of Type I and Type II errors, a Type I error in this situation would occur when the model predicts that an RR firm is a NRR firm, while a Type II error occurs when the model predicts that a NRR firm is an RR firm. Generally, Type I and Type II errors are not equally costly. In many situations, specific group classification rates are of relative interest and importance (Bernardi and Zhang 1999). Revenue restatement prediction is one such situation.

Misclassification costs in the case of a revenue restatement prediction are dissimilar. Auditors are likely to be more concerned about Type I errors; they do not want to issue clean opinions for firms that deserve qualified audit reports. Financial analysts also are more concerned about RR predictions because revenue restatements tend to invoke significant negative stock price reactions. In light of the legal and other consequences associated with Type I errors (giving a clean opinion to a firm that should have a qualified report), auditors, portfolio managers, and investors are more likely to be concerned with minimizing Type I errors. Type II errors are similar to false positives, and financial statement users are likely to consider this to be less important. In this context, the superior predictive power of the neural network model in predicting RR cases assumes special significance and is a key contribution of this paper. Because the true cost of misclassification is not observable (Koh 1992), prior studies have used relative cost ratios to examine the magnitude of misclassification costs. The ratio of Type I to Type II error rates is used to estimate the relative cost ratio. Relative cost ratios of 1:1, 10:1, 20:1, 30:1, 40:1, and 50:1 were used (Etheridge et al. 2000). The estimated relative cost (*RC*) of each earnings restatement prediction model is computed using the following formula (Koh 1992):

$$RC = [PI \times CI] + [PII \times CII]$$

where:

PI = the Type I error rate;
CI = the relative cost of a Type I error;
PII = the Type II error rate; and
CII = the relative cost of a Type II error.

Type I error rate (*PI*) and Type II error rate (*PII*) are reported in Table 3 for all three models. A relative cost ratio of 10:1 implies that the Type I error is ten times as costly as the Type II error.

The performance rankings of the three RR prediction models (NN, MDA, and Logit) for the various cost ratios are reported in Table 4. To compare the performance of the three models, a simple "sum of ranks" measure for different relative costs is used. This rank-sum measure yields a score of 8 for the NN model, 11 for the MDA model, and 17 for the Logit model. Excluding the 1:1 cost ratio, the rank-sum measure yields a score of 5 for the NN model, 15 for the Logit model, and 10 for the MDA model. The rank-sum measure comparison suggests that the NN model performs best with the lowest relative costs, followed by the MDA model and then the Logit model.

SUMMARY

This paper has outlined the process of training and testing a neural network model to predict revenue restatements. The purpose of this study was to develop and test a prototype neural network model that evaluates whether or not a firm is likely to restate its revenues (and consequently earnings). An initial sample was used to train the neural network, and a separate holdout sample was then used to validate the neural network predictions. The performance of the neural network was also compared with two conventional statistical models, a Multiple Discriminant Analysis model and a Logit model. The results show that the neural network model has superior predictive power for predicting revenue restatement firms when compared to the MDA and Logit models. The Logit and MDA models predict nonrevenue restatement firms better. Moreover, when misclassification costs are included, the neural network model performs the best with the lowest relative misclassification cost.

TABLE 4
Relative Costs

Cost Ratio	NN	MDA	Logit
Panel A: Estimated Relative Costs by Model			
1:1	0.407	0.316	0.355
10:1	0.205	0.385	0.424
20:1	0.184	0.392	0.432
30:1	0.176	0.394	0.434
40:1	0.172	0.396	0.436
50:1	0.170	0.397	0.437
Average	0.219	0.380	0.419
Panel B: Model Rank by Estimated Relative Cost			
1:1	3	1	2
10:1	1	2	3
20:1	1	2	3
30:1	1	2	3
40:1	1	2	3
50:1	1	2	3
Rank Sum	8	11	17

Greenstein and Welsh (1996) suggest that Logit and neural network models are complementary techniques and should not be viewed as alternatives. All three of the models tested in this study should be of interest to auditors and financial analysts.

The results of this study are consistent with Fanning and Cogger (1998), who concluded that there is the potential to detect fraudulent financial statements using neural networks trained on publicly available information. Green and Choi (1997) show that neural networks have excellent potential as a fraud investigation and detection tool given their aggregate error rate of only 25 percent. They suggest that one of the most significant benefits of neural networks is their ability to simultaneously evaluate all data input, whereas as traditional models used by auditors tend to evaluate data in isolation. The results of this paper on detecting revenue restatements combined with the results of other studies that have used neural networks to study fraud, bankruptcy prediction, and financial health suggest that the neural networks are a promising tool for accounting and finance practitioners to use when evaluating businesses.

In light of the well-publicized lawsuits against auditors, a neural network model to improve audit quality by predicting which firms are likely to restate their revenue (and earnings) is an important contribution to the field. A model that predicts revenue restatements would also be of value to investment analysts and portfolio managers who make earnings predictions and investment decisions based on current earnings forecasts. A prediction model is especially valuable in the current environment of more limited company disclosures due to the requirements of Regulation FD.

Limitations

Although neural networks are an important tool to use in developing prediction models, they do suffer from a few key limitations that impair their ability to produce highly accurate predictions. Most of the firms used in this study are large firms. Therefore, the results may not hold for firms of all sizes. Due to data limitations, only one year's data was used. Therefore, the results may be valid only for this set of data and training conditions. Another data limitation is the fact that only financial ratios and one governance variable were used in this study to develop the neural network model. In the real world, auditors and analysts may have numerous other pieces of information that may impact the decision on earnings restatement. For example, Brown and Caylor (2004) study 51 factors that encompass eight corporate governance categories and find eight factors that are most often associated with good performance, and seven factors that are most often associated with poor performance. In future research, we hope to examine whether firms with poor corporate governance characteristics are more likely to use inappropriate revenue recognition practices.

In addition to the data limitations, model limitations also exist. For example, there is no widely accepted theory to guide network topology. Various decisions such as the number of hidden layers, number of nodes, training tolerance, and testing tolerance must be made by trial and error. Since training is significantly influenced by the chosen parameters, there could be better neural network solutions that have not been explored with the same data. It should be pointed out that this neural network model does not replace the need for careful judgment on the part of professionals. Neural networks operate as "black boxes" and do not provide information about the statistical significance of independent variables like statistical models do. Future neural network research in this area could utilize techniques that provide more transparency regarding the relative predictive ability of the independent variables included in the study.

Despite these limitations, the results of this neural network model provide the first step toward gaining important insights that are useful for auditors, analysts, and investors who may be interested in identifying specific contexts in which revenues are likely to be manipulated. In future research, we hope to employ a technique, such as the modified genetic algorithm approach (Sexton et al. 2003) or the gradient descent procedure (Kruschke 1989), that will yield more generalizable results and identify the attributes that separate firms that restate revenues from firms that do not restate.

REFERENCES

Abdolmohammadi, M. J., and W. J. Read. 2004. Are audit opinion modifications associated with future financial restatements? Paper presented at the 2004 Midyear Meeting of the Auditing Section of the American Accounting Association.

Altamuro, J., A. Beatty, and J. Weber. 2005. The effects of accelerated revenue recognition on earnings management and earnings informativeness: Evidence from the SEC Staff Accounting Bulletin No. 101. *The Accounting Review* 80 (2): 373–402.

American Institute of Certified Public Accountants (AICPA). 2002. Statement on Auditing Standards No. 99. *Consideration of Fraud in a Financial Statement Audit*. New York, NY: AICPA.

Barniv, R., A. Agarwal, and R. Leach. 1997. Predicting the outcome following bankruptcy filing: A three-state classification using neural networks. *International Journal of Intelligent Systems in Accounting, Finance and Management* 6: 177–194.

Bernardi, V. L., and G. P. Zhang. 1999. The effect of misclassification costs on neural network classifiers. *Decision Sciences* 30: 659–682.

Boritz, J., D. Kennedy, and A. Albuquerque. 1995. Predicting corporate failure: Using a neural network approach. *International Journal of Intelligent Systems in Accounting, Finance and Management* 4: 95–112.

Brown, C., and J. Coakley. 2000. Artificial neural networks in accounting and finance: Modeling issues. *International Journal of Intelligent Systems in Accounting, Finance & Management* 9: 119–144.

Brown, L. D., and M. L. Caylor. 2004. Corporate governance and firm performance. Working paper. Available at SSRN: http://ssrn.com/abstract=586423.

Byrne, J. 2002. Let's really clean up those numbers now. *Business Week* (July 15).

Callen, J. L., J. Livnat, and D. Segal. 2002. Accounting restatements: Are they always bad news? Working paper, University of Toronto.

Choi, S., and D. Jeter. 1992. The effects of qualified audit opinions on earnings response coefficients. *Journal of Accounting and Economics* (June–September): 229–247.

Dechow, P., R. Sloan, and A. Sweeney. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the sec. *Contemporary Accounting Research* 13: 1–36.

DeFond, M., and J. Jiambalvo. 1991. Incidence and circumstances of accounting errors. *The Accounting Review* 66 (3): 643–655.

Etheridge, H. L., and R. S. Sriram. 1997. A comparison of the relative costs of financial distress models: Artificial neural networks, logit and multivariate discriminant analysis. *International Journal of Intelligent Systems in Accounting, Finance and Management* 6: 235–248.

———, ———, and K. Hsu. 2000. A comparison of selected artificial neural networks that help auditors evaluate client financial viability. *Decision Sciences* 31: 531–550.

Fadlalla, A., and C. H. Lin. 2001. An analysis of applications of neural networks in finance. *Interfaces* 31 (4): 112–122.

Fanning, K., K. Cogger, and S. Srivastava. (1995). Detection of management fraud: A neural network approach. *International Journal of Intelligent Systems in Accounting, Finance and Management* 4: 113–126.

———, and ———. 1998. Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems in Accounting, Finance and Management* 7 (1): 21–41.

General Accounting Office (GAO). 2002. *Financial Statement Restatements: Trends, Market Impacts, Regulatory Responses, and Remaining Challenges*. USGAO-03-138 (October).
———. 2003. *Financial Statement Restatement Database.* USGAO-03-395R (January).
Green, B., and T. Choi. 1997. Assessing the risk of management fraud through neural network technology. *Auditing: A Journal of Practice & Theory* 16 (1): 14–28.
Greenstein, M. M., and M. J. Welsh. 1996. Bankruptcy prediction using *ex ante* neural networks and realistically proportioned testing sets. Paper presented at the Annual Meeting of the American Accounting Association.
Guerra, J. 2004. The Sarbanes-Oxley Act and the Evolution of Corporate Governance. *The CPA Journal* 74 (4): 14–16.
Hardgrave, B., R. Wilson, and K. Walstrom. 1994. Predicting graduate student success: A comparison of neural networks and traditional techniques. *Computers and Operations Research* 21 (3): 249–263.
Kinney, W., and L. McDaniel. 1989. Characteristics of firms correcting previously reported quarterly earnings. *Journal of Accounting and Economics* 11: 71–93.
Koh, H. C. 1992. The sensitivity of optimal cutoff points to misclassification costs of Type I and Type II errors in the going-concern prediction context. *Journal of Business Finance and Accounting* 19 (2): 187–197.
Kruschke, J. K. 1989. Distributed bottlenecks for improved generalization in back-propagation networks. *International Journal of Neural Networks Research and Applications* 1 (1): 187–193.
Lawrence, J., and J. Frederickson. 1993. *Brainmaker: User's Guide and Reference Manual*, Nevada City, CA: California Scientific Software Press.
———. 1994. *Introduction to Neural Networks: Design, Theory, and Applications*, Nevada City, CA: California Scientific Software Press.
Lee, B. B., and B. Choi. 2002. Company size, auditor type and earnings management. *Journal of Forensic Accounting* 3: 27–50.
National Commission on Fraudulent Financial Reporting (NCFFR) 1987. *The Report of the National Commission on Fraudulent Financial Reporting.* Available at: http://www.coso.org/Publications/NCFFR.pdf.
———. 1999. *The Committee of Sponsoring Organizations of the Treadway Commission's Report on Fraudulent Financial Reporting 1987 to 1997.* Available at: http://www.coso.org/publications/FFR_1987_1997.PDF.
O'Leary, D. E. 1987. Validation of expert systems—With applications to auditing and accounting expert systems. *Decision Sciences*, 18: 468–486.
———. 1998. Using neural networks to predict corporate failure. *International Journal of Intelligent Systems in Accounting, Finance and Management* 7: 187–197.
Owers, J., R. Rogers, and C. Lin. 2002. The information content of different categories of earnings restatements. *International Business and Economics Research Journal* 1 (5): 71–83.
Palmrose, Z-V., V. Richardson, and S. Scholz. 2004. Determinants of market reactions to restatement announcements. *Journal of Accounting and Economics* 37: 59–89.
———, and S. Scholz. 2004. The accounting causes and legal consequences of non-GAAP reporting: Evidence from restatements. *Contemporary Accounting Research* 21 (1): 1–41.
Ragothaman, S. 2001. A neural networks approach to predicting going concern audit reports. *New Review of Applied Expert Systems and Emerging Technologies* 7: 33–50.
———. 2003. A neural networks approach to predicting corporate illegal behavior. *Journal of Forensic Accounting* 4 (2): 181–198.
Richardson, S., I. Tuna, and M. Wu. 2002. Predicting earnings management: The case of earnings restatements. Working paper, October 2002. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=338681.
Schipper, K. 1989. Commentary on earnings management. *Accounting Horizons* 3 (4): 91–102.
Securities and Exchange Commission (SEC). 1999. *Revenue Recognition*. Staff Accounting Bulletin No. 101. Washington, D.C.: Government Printing Office.

———. 2000. *SEC Accounting and Auditing Enforcement Release No. 1350*. Available at: www.sec.gov.

Sexton, R. S., R. S. Sriram, and H. Etheridge. 2003. Improving decision effectiveness of artificial neural networks: A modified genetic algorithm approach. *Decision Sciences* 34: 421–442.

Shawver, T. J. 2005. Merger premium predictions using a neural network approach. *Journal of Emerging Technologies in Accounting* 2: 61–72.

Stallworth, L., and D. Digregorio. 2004. Improper revenue recognition. *Internal Auditor* (June): 53–56.

Turner, L., R. Dietrich, K. Anderson, and A. Bailey. 2002. Accounting Restatements. Working paper, SEC and Ohio State University.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Wilson, R. L. 1992. Business implementation issues for neural networks. *Journal of Computer Information Systems* 32 (Spring): 15–19.

———, and R. Sharda. 1994. Bankruptcy prediction using neural networks. *Decision Support Systems* 11: 545–557.

Wu, M. 2002. Earnings restatements: A capital market perspective. Working paper, New York University.

Xie, B., W. N. Davidson, and P. DaDalt. 2003. Earnings management and corporate governance: the role of the board and the audit committee. *Journal of Corporate Finance* 9: 295–316.