A Data Analytic Approach to Predicting Firms that Corrected Prior Period Misstatements

Srinivasan Ragothaman
The University of South Dakota

Hyun Woong (Daniel) Chang
University of North Texas

Karen J. Davies
The University of South Dakota

When Staff Accounting Bulletin 108 (SAB 108) was issued by the Financial Accounting Standards Board (FASB) in late 2006, US companies were obligated to use the dual approach – the "rollover" and the "iron curtain" approaches resulting in the correction of the prior period errors in both the income statement and the balance sheet. We call these SAB 108 adopters PPA (prior period adjustment) firms. We present a comparative examination of three data analytical models (decision tree, discriminant analysis, and logistic regression) in predicting PPA firms. While analyzing the holdout sample, the decision tree model is more accurate in predicting PPA firms than the other two. Our recommendation is to apply all three models to predict PPA firms and then use the assessment provided by two or more models.

INTRODUCTION

The U.S. Securities and Exchange Commission (SEC 2006) provides guidance with respect to correcting errors that have built up over time in the books. This recent guidance is contained in Staff Accounting Bulletin 108 (SAB 108) issued by the SEC in 2006 and it explains the process to decide whether cumulative errors are material and if so, how to deal with them. Before 2006, companies used to consider qualitative and quantitative factors to decide if the uncorrected errors are material and dealt with them in a variety of ways. The two principal approaches to correct cumulative errors are the rollover approach and the iron curtain approach. In the past, companies have used either the iron curtain or the rollover approach, correcting the error in either the income statement or in the balance sheet (Keune and Johnstone, 2015). For example, if a company overstated the interest payable for $400 each year for the past 5 years, that company would have a total of $2,000 that has been misstated. Under the iron curtain approach, the entire $2,000 would be corrected by the company by decreasing interest payable in the current year by $2,000. In contrast, the company would only adjust the current year's misstatement of $400 under the rollover approach. Using either the iron curtain approach or the rollover approach has consequences for the financial statements of the company. If the iron curtain approach is used, the company's income statement would be misstated by $2,000 for the current year due to corrective adjustments, while the balance sheet would be correct. On the other hand, if the company uses the rollover approach, the income statement would be correct, but the balance sheet would be misstated by $1,600. Since the previous years' misstatement does not get corrected under the rollover approach, this misstatement would continue to remain on the books. If there are more misstatements in the future, such mistakes could accumulate and significantly distort the company's balance sheet.

After SAB 108 (which went into effect in late 2006) companies were compelled to use what is known as the dual approach to correct the cumulative errors, if the error is considered material. The dual approach is a compromise between the two methods, requiring firms to not only correct the current misstatement, like the rollover approach, but also correct for cumulative effects of previous misstatements, like the iron curtain approach. For the aforementioned example, the company would adjust the current misstatement of $400, like in the rollover approach. In addition, the company would have to adjust for the effects of previous misstatements by misstating the current year's income statement by $1,600. As result, the income statement is misstated by a lesser amount, while the balance sheet would be no longer misstated.

Prior period errors originate from multiple sources – arithmetic errors, GAAP errors, improper revenue recognition, aggressive liability estimates and others. The following questions will be investigated in this study: Is there a difference in auditors (Big 4 vs smaller auditors) between SAB 108 firms and control firms? Are more clients of smaller auditing firms correcting errors under SAB 108? Are SAB 108 firms less profitable? Are SAB 108 firms smaller? Do SAB 108 firms have larger Tobin's Q? Are SAB 108 firms more or less efficient in utilizing total assets? The purpose of this paper is to analyze the financial characteristics of companies that adopted SAB 108**.** The research study addresses the following two interrelated objectives: 1) To capture relevant patterns in the financial data structures of SAB 108 adopters using advanced statistical techniques such as Discriminant analysis, Logistic regression (logit) and Decision Tree and 2) To compare the relative predictive performance of these three models when using a Holdout sample.

Section 2 provides a brief summary of prior research. Section 3 describes the sample process and the data collected for analysis. Section 4 describes the logistic regression methodology and the results of both univariate and multivariate analyses. The last section provides a brief summary of the research study.

PRIOR RESEARCH

Since modern capital markets operate in a system where the owners (i.e. stockholders) are separate from management, it is important for the management to provide accurate information to the investors (Healy and Palepu, 2001). Among many pieces of information disclosed by management, financial statements are arguably the most important pieces of information provided to the investors. However, a problem arises when this information is incorrect, since investors' judgement and decisions were based on inaccurate information.

When examining SAB 108 misstatements specifically, Omer, Shelley, and Thompson (2012) find that firms that disclose SAB 108 misstatement had negative returns around the time of the disclosure. Such results suggest that investors are sensitive to misstatements, even if these misstatements relate to SAB 108, and thus were likely immaterial in the previous periods (Omer et al., 2012). Investors also react negatively towards misstatements that are not related to SAB 108. When these misstatements are severe and a restatement must be issued, industry peers, along with the restating firm, face lower stock prices (Gleason, Jenkins, & Johnson, 2008). Palmrose, Richardson, and Scholz (2004) find that the certain misstatement factors, such as existence of fraud, number of accounts affected, and reduction in income, affect the magnitude of negative returns for the misstating firms. Thus, better understanding misstatements will be beneficial to issuing firms' stakeholders, such as investors, regulators, and auditors.

Since misstatements impact many related parties, understanding which firms are more likely to misstate or restate is also important. Prior studies examine the nature of firms that misstate or restate their financial statements. Srinivasan, Wahid, and Yu (2015) find that foreign companies, especially those from countries that have weaker rule of law, that are listed in the U.S. are less likely to restate than their U.S.-based counterparts. The authors note that while more frequent restatement may make U.S. firms seem to have more problematic financial statements, because comparable foreign firms' accounting quality is worse than those of U.S. firms, it is more likely that foreign firms are not restating inaccurate financial statements (Srinivasan et al., 2015). Thus, when considering likelihood of misstatements, considering national characteristics could be useful.

Firm-specific measures can also be used to examine which firms are more likely to misstate. Dechow, Ge, Larson, and Sloan (2011) find that accrual quality and performance measures –financial and non-financial – are low when firms are misstating. They also find that firms tend to have more financing and off-the-books activities when the firms misstate (Dechow et al., 2011). These characteristics can serve as warning signs to the firm's stakeholders regarding the likelihood of an earnings misstatement, which can lead to earlier resolution of the problem. Also, since misstatements that are not fixed in time can accumulate and lead to restatements in the future, these factors can be useful in predicting restatements. A firm's earnings management tendencies, as measured by financial statement bloat, is another factor that determines the likelihood of misstatements (Ettredge, Scholz, Smith, and Sun, 2010). In their research, Ettredge et al. (2010) also state that if earnings management is indeed one of the causes of misstatements, proxy variables for earnings management could be also used to measure misstatement likelihood.

Characteristics of a firm's internal audit committee can affect a firm's likelihood of misstatement. Keune and Johnstone (2015) find that audit committees that have higher levels of short-term stock option compensation tend to waive more misstatements that may put the firm at a disadvantage in the short-term. Conversely, audit committees that have higher levels of long-term stock option compensation tend to waive more misstatements that could hurt the firm in the long-term (Keune and Johnstone, 2015). However, Audit committees that have members with greater financial expertise are less likely to permit managers to let material misstatements go uncorrected (Keune and Johnstone, 2012).

External auditors also play a role in a firm's misstatement. When the relationship between firms' restatements and external auditors' audit effort is examined, Lobo and Zhao (2013) find that the two have a positive relationship. While such results can be counterintuitive, since higher audit effort should lead to higher audit quality and fewer restatements, audit effort can be higher for firms that are already in financial trouble (Labo and Zhao, 2013). Thus, when controlled for other factors, such as auditors' risk adjustment, Lobo and Zhao (2013) find that audit effort and restatement likelihood have an inverse relationship. Also, as audit fee increases, auditors become much less likely to permit managers to not correct material misstatements (Keune and Johnston 2012). Thus, both internal and external audit related factors influence a firm's likelihood to misstate or restate financial statements.

Firms that issue multiple restatements can exhibit interesting characteristics. Files, Sharp, and Thomson (2014) examines the factors that are related to firms that issue multiple restatements in relatively short periods of time. They find that firms that issue repeat restatements are more likely to be not audited by Big 4 audit firms and that *ex ante* accounting quality is lower for the firms that issue multiple restatements. It should be noted that Lawrence, Minutti-Meza, and Zhang's (2011) suggest the differences in audit quality between Big 4 audit firms and non-Big 4 audit firms are attributable to client characteristics rather than the audit quality of the audit firms. Thus, Files et al.'s (2014) finding could suggest that the repeat restatements could be caused by the characteristics of the firms that are being audited by non-Big 4 audit firms, rather than the audit quality of the non-Big 4 firms themselves.

Keune and Johnstone (2009) suggest that there are distinguishing characteristics for companies that have SAB 108 adjustments. They find that firms in certain industries (banking, insurance, and real estate) are associated with more SAB 108 reports, which could be attributed to the fact that firms in this industry are more likely to be audited by smaller audit firms (Keune and Johnston, 2009). The smaller size of SAB 108 firms' auditors seem to parallel the findings of Files et al. (2014), since smaller auditors are related to more restatements in both studies. Keune and Johnstone (2009) also find that SAB 108 adjustments are made by firms that are larger. This finding is contrary to research that show smaller firms are more likely to restate (Kinney and McDaniel, 1989; DeFond and Jiambalvo, 1991).

Predicting misstatements is more difficult than understanding firm characteristics that are associated with misstating firms. Dechow et al. (2011) develop a probability measure that can show the likelihood of a firm misstating its financial statements. However, there is no specific measure in literature that predicts the likelihood of a firm issuing a SAB 108 disclosure statement. Thus, we explore various methods of predicting a firm's likelihood of issuing a SAB 108 disclosure due to uncorrected prior period misstatements. While Dechow et al. (2011) develop a likelihood measure for misstating firms in general, literature differences between characteristics of SAB 108 firms and general misstatement firms motivate us to develop a new measure for SAB 108 firms. Also, because Price, Sharp, and Wood's (2011) comparison between commercially developed measure and academically developed measure show academic measure's weaknesses, we compare traditional methods used in academia (i.e. logistic regression [logit] and discriminant analysis) with a more unconventional method (i.e. decision tree).

DATA DESCRIPTION

A sample of 119 firms was selected from a population of companies that had adopted SAB 108 in recent years and restated its financials. SAB 108 adopting firms were identified from the Audit Analytics database and these 119 companies adopted SAB 108 between 2006 and 2008 (Ragothaman 2013). Data from COMPUSTAT (Research Insight) for one year before the adoption on several operating and financial ratios such as total asset turnover, operating performance (return on equity), cash flow to sales ratio (a profitability measure), size (natural logarithm of sales), and inventory to total assets ratio (inventory intensity) were obtained for these firms. A control sample of another 119 firms (matched by industry) that had not adopted SAB 108 was randomly selected using the Yahoo Finance website. Financial ratios for the control sample were also obtained from the COMPUSTAT (Research Insight) data base. Because of missing data in COMPUSTAT, the final sample analyzed in this study consists of 82 SAB 108 adopters and 89 control firms. Support for using these specific variables is found in earlier research described in the prior research section. Table 1 provides a summary of descriptive statistics. For both SAB 108 firms and control firms, this table reports the mean, the standard deviation, and F-statistics for variables used in this study separately. Mean values for total asset turnover ratio and inventory to total asset ratios are higher for the SAB 108 firms when compared to control firms. However, the control firms have higher mean return on equity, profitability measures, and size measures than SAB 108 firms. The F-tests for equality of group means indicates that the SAB 108 firms are significantly smaller and have lower average returns on equity when compared to a set of control firms matched by industry. The F-test results also indicate that the mean total asset turnover ratio is significantly higher for the SAB 108 firms when compared to the control firms. The SIC industry codes were available for only 86 firms who adopted SAB 108. Of these, 16 companies belonged to business services industry, followed by 5 companies each in several industries including health services, communications, holding/investment services, electronic & electric equipment, industrial machinery and equipment, and food and kindred products.

STATISTICAL RESULTS

TABLE 1
DESCRIPTIVE STATISTICS

Variables	Firm Code	N	Mean	Std. Deviation	F-statistic
ROE	1	82	8.437	23.852	-2.730 [c]
	0	89	18.515	50.260	
TATO	1	82	1.148	0.882	3.109 [c]
	0	89	0.921	0.812	
LNSALES	1	82	6.799	1.743	-28.375 [a]
	0	89	8.288	1.901	
CFSALES	1	82	0.101	0.097	0.324
	0	89	0.117	0.222	
INVTA	1	82	0.092	0.110	1.699
	0	89	0.071	0.095	

Firm code: 1 = SAB 108 Firm; 0 = Control Firm
[a] two-tailed significance at < 0.01 level
[b] two-tailed significance at < 0.05 level
[c] two-tailed significance at < 0.10 level

A summary of Pearson correlation coefficients for the explanatory variables is provided in Table 2. There are not many strong correlations among the independent variables. Total asset turnover ratio is positively correlated with cash flow to sales ratio (a profitability measure). Return on equity is positively correlated with size, cash flow to sales, and total asset turnover ratio. There is a strong positive relationship between inventory intensity and total asset turnover ratio. Size (logarithm of sales) is positively associated with inventory to total assets ratio. Cash flow to sales ratio and inventory intensity are negatively related. Even though some of these relationships among independent variables are significant at conventional levels, none of the correlations are greater than 0.270 except one. There is one large correlation at 0.507 among 10 correlations among independent variables. According to Judge, Griffiths, Hill and Lee (1985), multicollinearity problems arise only when the correlations among explanatory variables are higher than 0.8. Hence, the degree of collinearity present among independent variables appears to be too small to invalidate estimation results. Also the variance inflation factor (VIF) values are all between 1.100 and 1.441 indicating no multicollinearity problem.

TABLE 2
PEARSON CORRELATION COEFFICIENTS

	ROE	TATO	LNSALES	CFSALES	INVTA
ROE	1.000				
TATO	0.085	1.000			
LNSALES	0.091	0.085	1.000		
CFSALES	0.270	-0.162	0.228	1.000	
INVTA	-0.018	0.507	0.178	-0.216	1.000

MULTIVARIATE TESTS – LOGISTIC REGRESSION (LOGIT)

Using the independent variables in a multivariate context, however, allows one to examine their relative explanatory power and can lead to better predictions since the information contained in the cross-correlations among variables is utilized. A primary objective of many multivariate statistical techniques is to classify entries correctly into mutually exclusive groups. Discriminant analysis, PROBIT, and LOGIT represent such multivariate models.

In this study, the following logistic regression (LOGIT) model is proposed:

$$\Pr (Y=1|X) = F (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_K x_k)$$

The dependent variable Y is a dichotomous (0, 1) variable representing the two groups, SAB 108 firms (Y=1) and control firms in the same industry (Y=0). The independent variables X_1, X_2, ... X_K include return on equity, total assets turnover ratio, size (logarithm of sales), cash flow to sales, and inventory intensity measures described in the previous section. Specifically these explanatory variables are:

ROE = Return on Equity
TATO = Total Asset Turnover
LNSALES = Natural logarithm of sales revenue
CFSALES = Cash flow to Sales ratio
INVTA = Inventory to Total Assets ratio

Decision trees are tree-based models that provide a useful way to predict the category (dependent variable) based on a set of predictor (independent) variables. The data are recursively partitioned into two groups based on the predictor variables. This process is repeated until the response (dependent) variable becomes homogeneous. The sequence of binary partitions of the predictor (independent) variables is displayed as a binary decision tree. The purpose of DA is to determine the linear combination of the explanatory variables that best discriminates between groups that are partitioned. DA is often applied to problems where the dependent variable is dichotomous, as in the current case. DA classifies entries into mutually exclusive groups by maximizing the inter-group to intra-group variance-covariance from a set of predictor variables.

RESULTS AND DISCUSSION

Table 3 describes the classification results of the five-variable DA (discriminant analysis), and Logit (logistic regression) models as well as the Decision Tree results. Training sample results are provided in Panel A of this table indicating that the Decision Tree (DT) model classifies 86 percent of the PPA firms and 54.9 percent of the NPPA cases correctly.

TABLE 3
TRAINING AND HOLDOUT SAMPLE RESULTS

	Classification Matrix		Predicted Group TOTAL	
Actual Group	**% Correct**	**PPA Firms**	**NPPA Firms**	**TOTAL**
Panel A: Training Sample Results				
Decision Tree				
PPA Firms	86.0	80	13	93
NPPA Firms	54.9	46	56	102
Type I Error: 14.0% Type II Error: 45.1%				
DA				
PPA Firms	76.8	63	19	82
NPPA Firms	69.7	27	62	89
Type I Error: 23.2% Type II Error: 30.3%				
Logit				
PPA Firms	70.7	58	24	82
NPPA Firms	74.2	23	66	89
Type I Error: 29.3% Type II Error: 25.8%				
Panel B: Holdout Sample Results				
Decision Tree				
PPA Firms	88.5	23	3	26
NPPA Firms	29.4	24	10	34
Type I Error: 11.5% Type II Error: 70.6%				
DA				
PPA Firms	76.9	20	6	26
NPPA Firms	65.5	10	19	29
Type I Error: 23.1% Type II Error: 34.5%				
Logit				
PPA Firms	76.9	20	6	26
NPPA Firms	79.3	6	23	29
Type I Error: 23.1% Type II Error: 20.7%				

Type I (II) error is defined as percentage of firms that were classified as NPPA (PPA) firms, while they were actually PPA (NPPA) firms. 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 decision tree model is to use different sets for development and testing. O'Leary also recommends validating AI models (machine learning algorithms, decision tree, and others) against other statistical models. Both validation methods are used to assess this neural network model's ability to classify firms into PPA and NPPA groups. Panel B shows that when the DT model is used to analyze the holdout sample, 88.5 percent of the PPA firms and 29.4 percent of the NPPA firms are grouped correctly.

A discriminant analysis (DA) model and a logit model are employed as content validation tools to evaluate the performance of the DT model. The discriminant analysis was performed using SPSS, and the results are described in Table 3. Panel A of Table 3 gives the training sample results indicating that the five-variable discriminant model classifies 76.8 percent of the PPA and 69.7 percent of the NPPA firms correctly. Panel B shows that when the discriminant model is used to analyze the holdout sample, 76.9 percent of the PPA and 65.5 percent of the NPPA firms are grouped correctly.

Much of the prior research in the area of financial data analytics provides a comparison of neural network results with a traditional linear technique such as DA (Brown and Coakley, 2000). We make additional comparisons in this paper against a non-linear 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. 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 DA, the fitted 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 70.7 percent of the PPA and 74.2 percent of the NPPA cases correctly. Panel B shows that when the holdout sample is analyzed using the Logit model, 67.9 percent of the PPA and 79.3 percent of the NPPA firms are grouped correctly. The holdout sample results indicate that while the DT is more accurate (88.5% accuracy) in predicting PPA firms, while the discriminant analysis and logit models both have an accuracy rate of 76.9 percent.

MISCLASSIFICATION COSTS

The comparison of the results of all three models shown in Table 3 suggests that the statistical models (DA and Logit) are more accurate in predicting the NPPA cases while the decision tree model is more accurate than the statistical models in predicting the PPA 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 a PPA firm is an NPPA firm while a Type II error occurs when the model predicts that an NPPA firm is a PPA firm. Typically, 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). PPA firm prediction is one such situation.

Misclassification costs in the case of a PPA 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 PPA predictions because prior period adjustments tend to invoke significant negative stock price reactions. In light of the stock price decline and other consequences associated with Type I errors (classifying a PPA firm as an NPPA firm), auditors, portfolio managers, and investors are more likely to be concerned with minimizing Type I errors. Type II errors are akin to false positives, and analysts and other users are likely to consider this to be less important. In this context, the superior predictive power of the decision tree model in predicting PPA cases is noteworthy and is a key contribution of this paper. According to Koh (1992), the true cost of misclassification is not observable. Hence, 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 by Etheridge, Sriram and Hsu (2000). The estimated relative cost (RC) of each PPA prediction model is computed using the following formula (Koh 1992)

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

where PI is the Type I error rate; CI is the relative cost of a Type I error; PII is the Type II error rate; and CII is the relative cost of a Type II error. Type I error rate (PI) and Type II error rate (PII) are reported in

Table 4 (Panel A) 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.

TABLE 4
RELATIVE COSTS

PANEL A: Estimated Relative Costs by Model

Cost Ratio	DT	DA	Logit
1:1	0.411	0.288	0.219
10:1	0.169	0.241	0.229
20:1	0.143	0.236	0.230
30:1	0.134	0.235	0.230
40:1	0.129	0.234	0.230
50:1	0.127	0.233	0.231

PANEL B: Model Rank by Estimated Relative Cost

Cost Ratio	DT	DA	Logit
1:1	3	2	1
10:1	1	3	2
20:1	1	3	2
30:1	1	3	2
40:1	1	3	2
50:1	1	3	2
Rank Sum	8	17	11

The performance rankings of the three PPA prediction models (DT, DA, and Logit) for the various cost ratios are reported in Table 4 (Panel B). 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 DT model, 17 for the DA model, and 11 for the Logit model. Excluding the 1:1 cost ratio, the rank-sum measure yields a score of 5 for the DT model, 10 for the Logit model, and 15 for the DA model. The rank-sum measure comparison suggests that the decision tree (DT) model performs best with the lowest relative costs followed by the Logit model and then, the DA model. A key limitation of this study is the small sample size. Therefore care should be exercised in generalizing the results of this study.

SUMMARY

Since SAB 108 went into effect in late 2006, US companies have to use the dual approach. i.e., both the rollover and the iron curtain approaches in order to correct cumulative misstatements from prior periods in both the income statement and the balance sheet. Stock prices are negatively affected and audit costs increase when misstatements are disclosed. Being able to predict a firm's likelihood of having a misstatement and issuing a SAB 108 disclosure statement will allow the stakeholders, such as investors, regulators and auditors to adjust their risk assessment of the firm. The accounting quality of the firm's domicile, the financial expertise of and profit-sharing plans associated with the internal audit committee, the firm's industry, and the audit fee, audit effort and size of the external auditor are all identified nonfinancial characteristics of SAB 108 adopters. Financial measurements relating to the firm's size, accrual quality, level of financing, performance measures and earnings management practices can also serve as a warning sign about the likelihood of an earnings misstatement. This paper analyzes the financial characteristics of companies that adopted SAB 108 (PPA firms) and corrected prior period errors and compares them to a set of control firms in the same industry.

Variables representing size, earnings management, and profitability selected for this study include the return on equity, total asset turnover, sales size, cash flow to sales ratio and inventory to total asset ratio. The descriptive statistical results indicate that SAB 108 adopters will have a higher mean value for total asset turnover and inventory to total asset ratios, and a lower average return on equity, profitability measures and size measures than companies in the same industry that had not adopted SAB 108. The

Pearson Correlation Coefficients show there are no significant multicollinearity problems, indicating the variables can be used as independent explanatory variables.

Using the five explanatory variables, three data analytical models are used in this study to determine their accuracy in predicting PPA firms. A decision tree model predicts if a firm is PPA or NPPA based on the independent predictor variables. Using a discriminant analysis model and a logistic regression model incorporate the possible cross-correlations among the variables and provides a validation tool to evaluate the performance of a decision-tree model.

Classifying a PPA firm as a NPPA firm (Type I error) is more costly to investors, external auditors and financial analysts than classifying a NPPA firm as a PPA firm (Type II error). Most stakeholders are interested in accurate predictions of PPA firms, while minimizing Type I errors. Misclassifying a NPPA firm as a PPA firm is not as costly. Using the holdout sample, this study finds that the decision tree model more accurately predicts PPA firms with the lowest relative cost. While the discriminate model and logit model provide more accurate predictions for NPPA firms, they also provide validation at a lower accuracy rate than the Decision Tree model for the classification of PPA firms. All three models can be used to predict PPA firms and the final decision can be made based on the assessment provided by two or more of the models. In other words, all three models are complementary techniques and may not be viewed as alternatives.

REFERENCES

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

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

Dechow, P., Ge, W., Larson, C., & Sloan, R. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research,* 28(1), 17-82. doi:10.1111/j.1911-3846.2010.01041.x

DeFond, M., & Jiambalvo, J. (1991). Incidence and circumstances of accounting Errors. *Accounting review*, 643-655.

Etheridge, H.L., Sriram, R.S., & Hsu, H.Y,K. (2000). A comparison of selected artificial neural networks that help auditors evaluate client financial viability. *Decision Sciences,* 31(2), 531-550.

Ettredge, M., Scholz, S., Smith, K. R., & Sun, L. (2010). How do restatements begin? Evidence of earnings management preceding restated financial reports. *Journal of Business Finance & Accounting*, 37(3/4), 332-355. doi:10.1111/j.1468-5957.2010.02199.x

Files, R., Sharp, N. Y., & Thompson, A. M. (2014). Empirical evidence on repeat restatements. *Accounting Horizons*, 28(1), 93-123. doi:10.2308/acch-50615

Gleason, C. A, Jenkins, N. T., & Johnson, W. B. (2008). The contagion effect of accounting restatements. *The Accounting Review*, 83(1), 83–110. https://doi.org/10.2308/accr.2008.83.1.83

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1–3), 405–440. https://doi.org/10.1016/S0165-4101(01)00018-0

Judge, G., W. E. Griffiths, R. C. Hill, H. Lütkepohl and T. Lee. (1985). The theory and practice of econometrics. (John Wiley and Sons, New York, NY).

Keune, M. B., & K. M. Johnstone. (2009). Inferences from the revelation of previously uncorrected accounting misstatements. *Accounting Horizons* 23(1), 19–53.

Keune, M. B., & Johnstone, K. M. (2009). Staff Accounting Bulletin No. 108 disclosures: Descriptive evidence from the revelation of accounting misstatements. *Accounting Horizons*, 23(1), 19-53.

Keune, M. B., & Johnstone, K. M. (2012). Materiality judgments and the resolution of detected misstatements: The role of managers, auditors, and audit committees. *The Accounting Review,* 87(5).

Keune, M. B., & Johnstone, K. M. (2015). Audit committee incentives and the resolution of detected misstatements. *Auditing: A Journal of Practice & Theory*, *34*(4), 109-137. doi:10.2308/ajpt-51080

Kinney Jr, W. R., & McDaniel, L. S. (1989). Characteristics of firms correcting previously reported quarterly earnings. *Journal of accounting and economics*, 11(1), 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.

Lawrence, A., Minutti-Meza, M., & Zhang, P. (2011). Can Big 4 versus non-Big 4 differences in audit-quality proxies be attributed to client characteristics? *Accounting Review*, 86(1), 259-286. doi:10.2308/accr.00000009

Lobo, G. J., & Zhao, Y. (2013). Relation between audit effort and financial report misstatements: Evidence from quarterly and annual restatements. *Accounting Review*, 88(4), 1385-1412. doi:10.2308/accr-50440

O'Leary, D.E. (1987). Validation of Expert Systems - with Applications to Auditing and Accounting Expert Systems. *Decision Sciences*, 18, 1987, 468-486.

Omer, T. C., Shelley, M. K., & Thompson, A. M. (2012). Investors' response to revelations of prior uncorrected misstatements. *Auditing: A Journal Of Practice & Theory*, 31(4), 167-192. doi:10.2308/ajpt-10303

Palmrose, Z. V., Richardson, V. J., & Scholz, S. (2004). Determinants of market reactions to restatement announcements. *Journal of Accounting and Economics*, 37(1), 59–89. https://doi.org/10.1016/j.jacceco.2003.06.003

Price, I. A., Sharp, N. Y., & Wood, D. A. (2011). Detecting and predicting accounting irregularities: A comparison of commercial and academic risk measures. *Accounting Horizons*, 25(4), 755-780. doi:10.2308/acch-50064

Ragothaman, S. (2013). Characteristics of US companies that adopted SAB 108 and corrected prior period misstatements: An empirical analysis. *AT Business Management Review,* 9(2), 85-89.

Srinivasan, S., Wahid, A. S., & Yu, G. (2015). Admitting mistakes: Home country effect on the reliability of restatement reporting. *Accounting Review*, 90(3), 1201-1240. doi:10.2308/accr-50887

U.S. Securities and Exchange Commission (2006). *Considering the effects of prior year misstatements when quantifying misstatements in current year financial statements*. Staff Accounting Bulletin No. 108. Retrieved from http://www.sec.gov/interps/account/sab108.pdf