December 2018

Predicting Accounting Misstatements within Industries and Accounts

Emily Cokeley
Syracuse University

ABSTRACT

This study examines variables that may be useful in predicting accounting misstatements. Using a database of Accounting and Auditing Enforcement Release information and building on recent models and methodology, I separate the observations by industry to determine the firm and financial statement variables that are most useful in predicting the firms within specific industries that may have accounting misstatements. I also extend the previous models to determine the significant variables in predicting not only which *firms* may have misstatements, but also the *account(s)* in which a misstatement is likely to have occurred. These models use information that is readily available in the financial statements, making them useful to auditors, regulators, and other users of financial statements. Finally, I examined the consistency of the predictive variables over several time periods.

My findings suggest that several variables that were found to be significant in a generalized model in previous literature lack significance in more specialized models and that some variables that were found to have no significance in a generalized model in previous literature do have significance in more specialized models. Specifically, the variables "soft assets" and "issue" appear to be the most consistent predictors of misstatements across industries, accounts, and time.

Keywords: accounting misstatements; fraud prediction; AAER; misstating industries; misstated accounts

PREDICTING ACCOUNTING MISSTATEMENTS WITHIN INDUSTRIES AND ACCOUNTS

By

Emily D. Cokeley

B.S., State University of New York, College at Brockport, 2006
B.S., Bethany College, 2003

Dissertation
Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Business Administration

Syracuse University
December 2018

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my dissertation advisor, Dr. Susan M. Albring. Her insight, guidance, and encouragement throughout the dissertation process exemplified her dedication to my success. I am also immensely grateful to committee member and former dissertation advisor, Dr. Randal J. Elder (currently at UNC Greensboro, and SU Professor Emeritus), for sharing his time, insight, and in-depth knowledge of auditing and misstatements. Susan and Randy exercised extreme patience and I cannot express how glad I am that they persisted with me to see this dissertation through to its completion. I would also like to thank the members of my dissertation committee: Dr. Mark S. Beasley (NC State University), Dr. Kofi Appiah Okyere, and Dr. Donald Harter, as well as my oral chair, Dr. Kira Reed. I truly appreciate their time, insight, and suggestions to help bring this dissertation to its present form.

My thanks as well to my colleagues in the Ph.D. program for their support, especially Rapheal Joseph Hamilton, Willie Dion Reddic, Xiaolu Xu, Alfred Yebba, Ravindra Ramchandra Bangar, Chao Chen, Michael Hyman, David Witesman, and Omar Watts. I am grateful for their assistance throughout my time in the program and for their contributions to my learning.

Finally, without the support and sacrifices made by my family, completion of this program would not have been possible. My husband Kevin Curran and sons Andrew and Alexander have my endless appreciation not only for their support, but also for enduring my divided attention as I focused on completing this program. My thanks as well to my parents Jeff and Sue Cokeley for their encouragement throughout this process, and especially for my mother's generosity, spending time with the boys while I wrote this dissertation.

TABLE OF CONTENTS

LISTING OF TABLES

1. INTRODUCTION

Throughout the last two decades, accounting researchers have attempted to develop methods to predict and detect financial statement misstatements due to error or fraud. Seminal research on this topic has found that using simple financial ratios can allow interested financial statement users to predict misstating firms with some accuracy (Beneish, 1999; Bayley and Taylor, 2007). Further research has found financial variables that are significant in predicting misstating firms, although the calculation of the variables is more complex (Dechow, Ge, Larson, and Sloan, 2011). Both approaches to the prediction of misstating firms involve the creation and use of a scoring system to assign the likelihood that an individual firm may be misstating its financial statements. In addition, both approaches address the prediction of misstating firms, but do not extend further to more specialized prediction models that include a specification of the firms' industry, nor do they extend further to identify the individual account(s) that may be misstated.

To date, no research has attempted to create a more accurate measurement of the likelihood of misstatement beyond an overall assessment at the firm level. There is a lack of subsequent research designed to provide greater ability to predict misstating firms by drilling down by industry, and similarly, there is a lack of research which aims to predict which account(s) may be misstated by firms identified as likely to misstate.

Beasley, Carcello, Hermanson, and Neal (2010) find that misstatements are clustered in certain industries, indicating the importance of considering the industry in which a firm is operating. Consideration of industry risk and account-level risk is noted specifically in PCAOB Accounting Standard 2110 (AS 2110). Paragraphs 04 through 09 of AS 2110 note that the industry in which a firm is operating should be taken into consideration when assessing risk and that the auditor should obtain an understanding of the firm and the industry and environment in which it operates. Specifically, paragraph 05 notes that "risks of material misstatement can arise from a variety of sources," including the firm's industry and environment. Paragraphs 59 through 64 of AS 2110 go into further detail about the auditor's requirement to assess the risk of material misstatement, noting in particular the need to assess account- and assertion-level risks. AS 2110 requires that the auditor identify significant accounts and the possible sources of potential significant misstatements within each identified account.

Since recent scoring models provide a reasonable level of accuracy in predicting misstatements, it follows that using similar techniques to create scoring models specific to certain industries and individual accounts can potentially result in models with even *greater* power for misstatement prediction. This study attempts to determine the variables significant to the prediction of misstating firms by industry, and also determine the variables significant to the prediction of individual misstated accounts, thus filling the gap in the literature while adding to the information available to investors, creditors, regulators, and auditors.

To identify the variables significant in the prediction of misstating firms within industries and accounts, I follow the methodology used in Dechow et al. (2011). I first obtain Accounting and Auditing Enforcement Release (AAER) data, described fully in Dechow et al. (2011). The dataset that I work with for this paper expands the data timeframe (1980 to 2002) from that of Dechow et al. (2011) to include more current evidence on misstatements (1980 to 2008).

Next, I determine the variables that are likely to have predictive value for each of the selected industries and accounts. Because these variables are based on publicly available financial statement data, this increases the accessibility to financial statement users. In particular, this study focuses on predicting the likelihood of misstatement in the most frequently misstated industries and accounts. In studies performed for the Committee of Sponsoring Organizations of the Treadway Commission (COSO), Beasley, Carcello, and Hermanson (1999) and Beasley, Carcello, Hermanson, and Neal (2010) note that during both the period 1987 through 1997 and the period 1998 through 2007, the industries with the greatest frequency of misstated financial statements included computer hardware/software, other manufacturing, financial service industries, healthcare and health products, retailers and wholesalers, and other service providers. In addition, Beasley et al. (1999) and Beasley et al. (2010) find that the most frequently misstated accounts for the time periods 1987 through 1997 and 1998 through 2007 remained the same: revenue, inventory, accounts receivable, property, plant, and equipment, and liabilities and expenses.[1] In addition, an analysis of the AAER data showed that the computers, durable manufacturers, services, and retail industries were the top four most frequently misstated industries during the period 1980 to 2008. All other industries had less than 60 misstated firm years across the 29-year period studied in this paper. Accordingly, this study focuses on the industries and accounts that are most likely to involve materially misstated financial statements. An analysis of the selected industries and accounts will offer the greatest value to those assessing the information risk within the financial statements. As other industries and accounts are much less represented as having misstatements in the financial statements, analysis on these remaining industries and accounts would offer minimal incremental value.

Using logistic regressions based on misstating firms within these selected industries and logistic regressions based on the most frequently misstated accounts, I determine the variables that are significant to misstatement prediction within these industries and accounts. Based on a sample of 1,297 firms that were issued at least one AAER between 1982 and 2012, I find that the percentage change in soft assets (soft assets are defined as total assets less property, plant, and

[1] The COSO report does not break out specific liability or expense accounts.

equipment and cash) and the issuance of securities during the fiscal year is consistently significant in the prediction of misstatements within industries and accounts. In addition, I find that leverage and the existence of operating leases are significant in the prediction of misstating firms within the computer hardware and software industry, where these variables were not significant in a generalized model in previous literature. For the retail industry, change in same store sales, a newly tested variable, is found to be significant. For the service industry, I find no variables to be significant that were not significant in a generalized model, although some variables significant in a generalized model were not significant in the service-specific model.

I also find that several variables that were previously found to be useful in a general misstatement prediction model do not hold their significance when utilized in models of predicting misstatement by industry. For the computer hardware and software industry, these variables include rsst accrual, change in inventory, percent change in cash sales, percent change in return on assets, and abnormal change in employees.[2] For the retail industry, these variables include rsst accrual, percent change in inventory, percent change in cash sales, percent change in return on assets, abnormal change in employees, and change in operating leases. For the service industry, these variables include rsst accrual, percent change in receivables, percent change in return on assets, abnormal change in employees, and change in operating leases.

With regard to prediction models for misstated accounts, I find that change in inventory, soft assets, the existence of a lease, and the issuance of securities during the year were significant predictors of misstatements in all three studied accounts. In addition, I find that the change in receivables and change in operating leases were predictors of misstatements in the revenue and accounts receivable accounts, but not for the inventory account.

---

[2] See Appendix A for a definition of all variables.

To test the stability of the models, I use the prediction models over subperiods of the full sample. Many predictors of misstatements tend to remain consistent across time periods, whether for all industries and accounts or for industry- and account-specific prediction models.

This paper contributes to the accounting literature by being the first to determine the most relevant variables in the prediction of financial statement misstatements within specific industries and accounts. While other research has provided generalized misstatement prediction methodologies, none have drilled down to create predictions specific to industry or account. By utilizing industry- or account-specific predictive variables and models, creditors, investors, regulators, and auditors may be able to predict with greater power and accuracy which firms may be more likely to have misstated their financial statements and which accounts may be more likely to have been misstated. This study finds that some variables from a general model may be useful predictors in most industry- and account-specific misstatement models, while finding that other variables that were not previously included in or found significant in a general model do have predictive value in some industry-specific and account-specific models. These differences from a generalized model from prior literature indicate that greater predictive accuracy is possible through the use of variables tailored to each industry-specific and account-specific model.

The remainder of this paper continues as follows: Section 2 reviews prior literature, Section 3 discusses the motivation and hypotheses, Section 4 describes the data used in this study, Section 5 delineates the empirical results, and Section 6 concludes the paper and offers avenues for future research.

2. BACKGROUND

2.1 Fraud Risk Factors

Identifying factors that can help predict fraud has long been a topic of interest to auditing standard setters and academic researchers. For example, SAS 82 (AICPA 1997) provided auditors with examples of fraud risk factors that could be used to help predict fraud. Five years later, this standard was replaced by SAS 99 (AICPA 2002) which incorporates the fraud triangle of opportunities, incentives, and attitudes/rationalization.[3] Early research attempted to identify factors associated with fraud firms. For example, Beasley (1997) found that the proportion of independent board members is lower for firms that experience financial fraud compared to a matched sample of non-fraud firms. Farber (2005) and Abbott, Parker, and Peters (2004) similarly find that fraud firms have poor governance compared to non-fraud firms.

These studies look at corporate governance as one dimension of firms that engage in fraudulent financial reporting. In contrast, other studies use multivariate prediction models that attempt to comprehensively investigate the characteristics of firms that engage in financial reporting fraud.

2.2 Early Misstatement Prediction Models

Academic research has described several approaches to predict misstating or at-risk firms, beginning with Altman's Z-score (1968), continuing with Beneish's M-score (1999) and more recently, through Dechow et al.'s F-score (2011). Different approaches and data sources have been used throughout prior literature, but no previous study has broached the realm of

---

[3] The current AICPA standard is AU-C 240 and is largely consistent with SAS 99.

going beyond the prediction of misstating firms and into the prediction of misstating firms by industry or accounts that a firm may be misstating.

Altman (1968) wrote a seminal paper in this area, stemming from his observation that academicians and practitioners find differing analytical methods more useful in the analysis of financial performance of businesses. Practitioners preferred the use of ratio analysis while academicians preferred the use of statistical techniques. While his study's focus was to "bridge the gap" between statistical analysis and ratio analysis, it is convenient that Altman analyzed the prediction of business bankruptcy. In this study, Altman used a small, matched sample of firms: those that had declared bankruptcy and those that had not. Altman combined the use of ratio and statistical analyses by using multiple discriminant analysis (MDA) to create a model to predict the likelihood that a firm declares bankruptcy. The use of MDA allowed Altman to analyze twenty-two financial ratios of interest at the same time to determine the financial ratios that, together, most accurately predicted whether a firm would declare bankruptcy. He found that considering the following ratios together resulted in a more accurate prediction model than considering several ratios independently: 1. Working capital/Total assets; 2. Retained earnings/Total assets; 3. Earnings before interest and taxes/Total assets; 4. Market value equity/Book value of total debt; and 5. Sales/Total assets.

The prediction model Altman arrived at considers the predictive value of each of these five ratios and weights them accordingly, using their coefficient estimates. By multiplying a firm's ratios by the corresponding coefficients and adding the results of each product, Altman arrived at a Z-score: the indicator of the probability of a firm declaring bankruptcy. He determined the Z-score above which about 95% of firms would declare bankruptcy as well as the Z-score below which no firm in the sample declared bankruptcy. He referred to the area between these two Z-scores as the “zone of ignorance,” or in other words, the prediction of bankruptcy for firms with a Z-score falling into this area was less certain. Altman also found that his prediction model could accurately predict the occurrence of bankruptcy up to two years prior to the declaration, with diminishing accuracy for longer periods. In summary, Altman’s Z-score model claimed 95% accuracy in the year prior to the bankruptcy and 72% accuracy two years prior to the bankruptcy. The Z-score model’s accuracy decreases to less than 50% when attempting to predict bankruptcy three to five years prior to the bankruptcy.

While Altman was not the first to use multiple discriminant analysis in the financial or business setting, he was instrumental in using MDA to combine two useful methods of analysis to create a prediction model that is generally accessible to both practitioners and academicians. Approaches similar to his methodology have been used in subsequent misstatement studies, notably Beneish (1999) and Dechow et al. (2011), as well as in this dissertation.

Beneish (1999) used indices of eight financial ratios to create a model to predict the likelihood that a firm manipulated its earnings. Beneish selected 49 firms identified by AAERs as earnings manipulators and an additional 25 firms identified by media outlets as earnings manipulators, for a total of 74 firms that had manipulated their earnings at some point between the years of 1982 and 1992. The manipulators were compared to 2,332 non-manipulator firms matched with the manipulator firms based on industry and year. For each firm, he calculated the following eight financial statement ratio indices by comparing year t to year t-1: 1. Days sales in receivables; 2. Gross margin; 3. Asset quality; 4. Sales growth; 5. Depreciation; 6. Sales, general, and administrative expenses; 7. Leverage; and 8. Total accruals to total assets. Using an unweighted probit estimate to determine the significance of these indices, he found that days sales in receivables, gross margin, asset quality, sales growth, and total accruals to total assets offered usefulness in the prediction of an earnings manipulation firm. Even though three of the eight tested variables were found to lack significance, Beneish included all of the variables in his prediction model. Like Altman (1968), Beneish added the products of the coefficient estimates and the corresponding ratios for each firm to arrive at the M-score. This resulted in about 50% accuracy in the prediction of earnings manipulation firms. Beneish found no major difference in the model's predictive accuracy resulting from the removal of the non-significant variables and adjusting the coefficient estimates.

The five significant variables found in Altman (1968) and the five significant variables found in Beneish (1999) have some similarities. Components of several of the variables include sales, total assets, current assets, current liabilities, and certain components of equity.

To determine the usefulness of ratio analysis as a method for identifying misstatements, Kaminski, Wetzel, and Guan (2004) tested twenty-one simple ratios, similar to those used in an auditor's analytical procedures, that can be derived directly from financial statements. The study selected 79 misstating firms as identified in AAERs issued between 1982 and 1999 (fiscal years analyzed spanned from 1975 to 1999) and matched the firms to similar, non-misstating firms. A longitudinal study of these twenty-one ratios was conducted, spanning three years prior to and three years subsequent to the fraud year. In univariate and multivariate analyses, sixteen of the twenty-one ratios were statistically significant for at least one of the seven years studied. However, only three of the ratios were significant for three consecutive years: fixed assets to total assets, total liabilities to total assets, and working capital to total assets. No ratios were consistently significant throughout the sample period.

Kaminski and Wetzel (2004) conducted a longitudinal examination of ten financial ratios on 30 matched-pair firms using chaos theory. None of the ratios exhibited stable behavior and they did not find any difference among the dynamics of these ratios for fraudulent and nonfraudulent firms.

A prediction model based on the significant ratios in Kaminski et al. (2004) resulted in misclassification of fraud firms between 58 and 98 percent of the time. Because of these results, the authors conclude that ratio analysis is not a reliable method of fraud firm identification. This study may have been impacted by a changing economic environment, as the study required a seven-year span for each misstating firm. The study stretched into the late 1990s, a time in which several major financial statement frauds were just starting to come to light. The effect of these frauds may have skewed the results of the ratio analysis in some fashion. Despite the conclusion by the authors, when used in conjunction with more complex ratios and/or accrual analysis, ratio analysis does appear to have some useful predictive validity (Altman 1968, Beneish 1999, Bayley and Taylor 2007, Dechow et al. 2011).

Bayley and Taylor (2007) note the same disconnect between practice and academia that Altman (1968) did, noting further that most earnings management literature subsequent to Jones (1991) follows the Jones or Modified Jones models, making incremental changes to the model in an attempt to find a more precise model of identifying overstated earnings. Bayley and Taylor believe that too much focus has been placed on avoiding Type I errors where the avoidance of Type II errors should be the greater focus. In normative terms, it is better to predict that a non-misstating firm has misstated their financial statements than to predict that a misstating firm has not misstated their financial statements.

Bayley and Taylor (2007) contend and find that using a simpler measure of accruals in conjunction with simple ratio analysis will have more power in identifying earnings manipulation than prior iterations of the Jones model. Using a sample of 129 earnings manipulators identified by AAERs as having manipulated their earnings between 1991 and 2003, the study matched each of these firms with five non-manipulators (the control group) and used logit analysis to determine the significant variables among six ratio indices. Bayley and Taylor use the same methodology as Altman (1968) and Beneish (1999) by multiplying the coefficient estimates by the corresponding accruals and ratios for each firm to arrive at an EM-score.

Univariate logistic regressions found that the operating accrual magnitude, sales index, accruals index, and inventory index were significant in the identification of earnings manipulation. The sales index is a measure of the reported net revenue compared to an estimate of unmanipulated net revenue. Unmanipulated net revenue is calculated using the change in the accounts receivable to net revenue ratio over time. The accruals index is a measure of earnings manipulation and is calculated using the current value of deflated operating accruals, lagged total assets, and the lagged current value of deflated operating accruals. The inventory index is a measure of earnings manipulation using inventory accounting techniques. This index is calculated using the current inventory to net revenue ratio and the lagged inventory to lagged net revenue ratio.

Multivariate logistic regressions found that operating accrual magnitude and the sales index were the only significant variables in earnings manipulation identification. Operating accrual magnitude is calculated as the sum of earnings before extraordinary items and depreciation and amortization expense, less cash flow from operations, scaled by total assets. The sales index is calculated by dividing net revenue by the estimate of non-manipulated net revenue. Bayley and Taylor (2007) compared the power of their EM-score model to the power of four other models based on the Jones model used in prior research (Dechow and Sweeney, 1995; Dechow et al., 2003), and found that the EM-score model has more power than the other four models, with about 65% classification accuracy. As a result, the authors recommended that future research in this field focus more on models that combine simple accrual calculations with other financial statement ratios while putting less focus on discretionary accrual models.

2.3 Non-Financial Measures in Fraud Prediction

The usefulness of non-financial measures as predictors of revenue fraud was studied by Brazel et al. (2009). The authors hypothesized that the difference between the year-over-year change in any non-financial measure and year over year change in revenue would be larger for fraud firms than for non-fraud firms. It was expected that the firms committing financial statement fraud would manipulate the financial data, but not the non-financial data that is made public, not always in the financial statements, but in other data sources. Non-financial data was hand-collected through searches of 10Ks, Proquest, Lexis/Nexis, Google, and other available resources. The authors identified their fraud sample using the AAERs in the Beasley et al. (1999) COSO report from 1987 to 1997, by conducting their own search of AAERs from 1998 to 2007, and by searching media sources for additional fraud cases. The final fraud sample consisted of 50 fraud firms, with fraud years ranging from 1994 to 2002. The fraud firms were each matched with one of their closest non-fraud competitors. Brazel et al. (2009) created a generalized variable to measure the difference between revenue growth (or reduction) and the growth (or reduction) of any non-financial measure (NFM):

$$\text{Capacity Diff}_t = \frac{(\text{Revenue}_t - \text{Revenue}_{t-1})}{\text{Revenue}_{t-1}} - \frac{(\text{NFM}_t - \text{NFM}_{t-1})}{\text{NFM}_{t-1}}$$

As an example, the study used employee growth as a single non-financial measure:

$$\text{Employee Diff}_t = \frac{(\text{Revenue}_t - \text{Revenue}_{t-1})}{\text{Revenue}_{t-1}} - \frac{(\text{Employees}_t - \text{Employees}_{t-1})}{\text{Employees}_{t-1}}$$

Using 18 control variables selected based on their use in prior literature (including leverage, Altman Z-score, and total accruals), the study calculated the difference between the means of each control variable as well as Capacity Diff and Employee Diff (Capacity Diff is calculated as the average of all of the non-financial measures that were available for each matched pair). In addition to a few control variables, the study found that both Capacity Diff and Employee Diff were significantly larger for the fraud firms than for the matched non-fraud firms, indicating that the Capacity Diff and Employee Diff both have discriminatory power in the detection of financial statement fraud. Finally, a logistic regression was run with Employee Diff as the independent variable of interest. Again, in addition to a few control variables, Employee Diff was significant to the detection of financial statement fraud.

The results of the Brazel et al. (2009) study indicate that similar to some financial measures and ratios, non-financial measures can be useful in the detection of financial statement fraud. This study was also able to establish benchmarks that can be used by those analyzing a firm's non-financial measures, which were generalized and not industry-specific, to determine the likelihood of financial statement fraud. The authors recognize that some of the firms committing financial statement fraud may in fact also fraudulently adjust the applicable non-financial measures, especially as time goes on, as the methods of fraud have a tendency to change over time. On a related note, many of the available non-financial data are firm-provided and not independently confirmed, which may have a negative effect on the reliability and usefulness of the non-financial measures. Finally, the authors question whether there may be a lead or a lag in the data, meaning that for example, a decrease in the number of employees may occur at the same time, before, or after revenue decreases. The introduction of non-financial measures into a misstatement prediction model introduces a slew of topics for future research, and more so as the non-financial data becomes more readily available.

Cohen et al. (2012) studied the availability of several non-financial measures gathered via corporate reporting of this data. Using a sample of ten firms in each of five industries, Cohen searched available data sources, including mandatory filings, websites, fact sheets, press releases, etc., for firm-provided information on six leading indicators of long-term value (future cash flows). In particular, Cohen et al. (2012) searched for information regarding market share, quality rankings, customer satisfaction, employee satisfaction, employee turnover, and innovation. The research found that there is great variability of reporting of these non-financial measures. Some firms may provide a significant amount of non-financial information while others provide very little. Whether a firm discloses certain non-financial information is also dependent on the size of the firm and the industry in which it operates.

The study found that of the six measures studied, the most frequently disclosed measures were market share and innovation, and the most frequently used non-financial disclosure sources were corporate websites and mandatory filings. This study highlights a potential selection bias: firms may choose to voluntarily disclose positive non-financial information while opting not to disclose negative non-financial information. In addition, the non-financial data provided by firms can be withdrawn from sources like corporate websites at any time. As a result, the authors note that a net benefit may result from the institution of mandatory disclosure and/or assurance by independent external auditors. The study included a sample size of 50 firms, and the results of this study indicate that finding reliable information reported on a consistent basis is unlikely. At this point in time, most non-financial information would need to be hand-collected, which is prohibitive, and the accuracy of the information found may not be reliable, as similarly noted by Brazel et al. (2009).

2.4 Recent Misstatement Prediction Models

Brazel et al. (2015) surveyed 194 nonprofessional investors who had traded individual stocks within the previous 12 months to determine relationships between investor perception of the prevalence of financial statement fraud, investor use of financial statement information, investors' conducting of their own fraud risk assessments, and investor use of red flags that could indicate fraud. The study found that the investors who use financial statement information and believe that financial statement fraud is prevalent also find it important to conduct their own fraud risk assessments. In addition, the study found that of those investors who conduct their own fraud risk assessments, they consider SEC investigations, pending litigation, violations of debt covenants, and management turnover as useful indicators of financial statement fraud. Three of the four noted red flags are non-financial measures, indicating that moderately sophisticated users of financial statements tend to find non-financial measures to be useful indicators of the risk of fraud, and make investing decisions based on their risk assessment. Brazel et al. (2015) also noted that some financial and non-financial measures were not considered important in the risk assessment decisions by the surveyed investors. Those measures include company size and age, need for external financing, and auditor quality (Big 4 vs non-Big 4 auditor).

Dechow et al. (2011) agree with the recommendation made by Bayley and Taylor (2007) to move toward models that combine a simple accrual calculation with other financial statement ratios, and created a model that includes the modified Jones model of discretionary accruals and performance-matched discretionary accruals, as well as working capital accruals. They also use a measure of accruals similar to working capital accrual, but included changes in long-term operating assets and long-term operating liabilities, as well as numerous financial and non-financial measures. Ultimately, Dechow et al. (2011) found that of the accruals variables tested, only the working capital accruals modified to include changes in long-term operating assets and long-term operating liabilities was significant in the prediction of misstating firms.

Dechow et al. (2011) reviewed all of the available AAERs (at that time) for companies cited for violating GAAP and created another scoring model for assessing the likelihood that any particular public company may be materially misstating its financial statements. The Dechow et al. (2011) prediction model (hereafter referred to as the Dechow model) uses information available in financial statements and from other sources (i.e., company websites for the number of employees and other nonfinancial measures, although it appears that the data for the Dechow model variables was obtained from COMPUSTAT and CRSP) in order to calculate specific variables used in the prediction model.

Using AAERs from 1982 to 2005, Dechow et al. (2011) collected information about firms that were found by the Securities and Exchange Commission (SEC) to have intentionally manipulated their annual financial statements. After determining the misstating firms based on AAER issuance, the Dechow et al. (2011) study analyzed several financial variables and ratios that were calculated using information taken from the financial statements, including two non-financial measures. The variables analyzed were broken down into the following categories (number of variables): accrual quality (9), performance (5), nonfinancial measures (2), off-balance-sheet information (4), and market-related incentives (8). The non-financial measures included were abnormal change in employees and abnormal change in order backlog. I note the abnormal change in employees variable in particular because the Dechow et al. (2011) study calculates this variable differently than Brazel et al. (2009). Dechow et al. (2011) calculates this variable as:

$$\text{ch\_emp} = \frac{(\text{Employees}_t - \text{Employees}_{t-1})}{\text{Employees}_{t-1}} - \frac{(\text{Total Assets}_t - \text{Total Assets}_{t-1})}{\text{Total Assets}_{t-1}}$$

No explanation was offered for their use of total assets, rather than the total revenue measure used in the Brazel et al. (2009) calculation. While the study did find the abnormal change in employees to be statistically significant, it is unclear if this variable's significance would increase or decrease if total revenues was used. Brazel et al. (2009) indicate that the number of employees is a measure of capacity for production of earnings, with the understanding that as capacity for production of earnings decreases, earnings (revenue) should decrease as well, and to a similar extent. The Dechow et al. (2011) calculation of this variable, using total assets instead of revenue, does not necessarily offer the same expected correlation. In addition to the number of employees, total assets is also a measure of capacity (as the number of production facilities decreases, so does the capacity for production of earnings, for example). The two measures both measure capacity for production for earnings, as opposed to measuring the *capacity* for production of earnings against *actual* reported earnings.

By running a series of logistic regressions on the variables, Dechow et al. (2011) found that 11 of the 28 tested variables were significant in the prediction of misstating firms; four in the accrual quality category; two in the performance category; one in the non-financial measures category; one in the off-balance-sheet category; and three in the market-related incentives category. The significance of these variables within their categories speaks to the opportunity and motivation involved when intentionally manipulating the financial statements; because accruals are subjective, they offer a greater opportunity to engage in earnings management, and meeting performance goals and a firm's participation in accessing the debt market provides the incentive behind earnings management. The significant variables are indicators providing insight into the likelihood that a firm has manipulated their financial statements. This finding corroborates the concept of the fraud triangle: that fraud is more likely to occur when there are pressures/motivation, opportunity, and rationalization present (Bell and Carcello, 2000; Rezaee, 2005; Hogan et al., 2008; Trompeter et al., 2013).

After determining the significant variables, the Dechow et al. (2011) study moved on to the second stage of their paper by developing a prediction model that assigns each significant variable (as determined in the first stage of their paper) a relative weight or value in predicting misstated financial statements. Combining those weights and the variable values for each firm-year, the model assigns each firm-year an F-score, indicating the probability that the financial statements associated with any particular firm-year are misstated.

The Dechow model uses an F-score of 1.4 as an indicator: over half of the misstated firms (as indicated by the AAERs) had an F-score of 1.4 or greater. The study also found that the F-scores for misstating firms increased in the years leading up to the misstated firm-year and decline after the misstatement firm-year.

The Dechow model appears to be reasonably accurate in predicting fraud firms (with accuracy of up to 69%), even taking into account the self-selection limitations imposed by the AAER issuance decisions made by the SEC (Aghghaleh, Mohamed, and Rahmat, 2016). How the SEC selects a firm to investigate for misstated financial statements is not public knowledge, although the investigation process may be initiated by firm-initiated restatements, media accusations or concerns about misstatements, or reports by whistleblowers. Because the SEC cannot and does not investigate and cite all firms that have manipulated their financial statements, there are an unknown number of firm years that are materially misstated but not reported as such through the issuance of an AAER. These firms and firm-years are included with the non-misstating firms in the data, instead of with the misstating firms, thus skewing the data to an unknown degree. A firm-year with an F-score below 1.4 may have actually manipulated the financial statements, but without being cited for such through an AAER, this firm would be included with the non-misstating firms. It is impossible to determine which firms fall into this problem area. Even though reliance on the SEC's investigations and AAERs can result in some errors in the data, AAERs are still considered to be a reasonable source for research data, as evidenced by their use by Beneish (1999), Brazel et al. (2009), Dechow et al. (2011), and many other researchers.

2.5 Prevalence of Misstatements

The Committee of Sponsoring Organizations of the Treadway Commission (COSO) commissioned a study of fraudulent financial statement occurrences, as cited by the SEC through AAERs. In the study titled "Fraudulent Financial Reporting: 1987 – 1997; An Analysis of U.S. Public Companies," Beasley et al. (1999) reviewed AAERs issued between January 1, 1987 and December 31, 1997, finding nearly 300 companies that were cited for the violation of Rule 10(b)-5 of the 1934 Securities Exchange Act or Section 17(a) of the 1933 Securities Act during that time frame. These two rules/sections were selected because they are the primary anti-fraud provisions related to financial statement reporting. Of these 300 firms, the authors randomly selected a sample of 204 firms to study further. The study examined specific characteristics pertaining to the company and to the management of the firms in the sample, finding trends and similarities (and in some instances, no trends or similarities) among the misstating firms. The findings of this study included details about the nature of the misstating firms, the nature of the control environments, the nature of the frauds, issues related to the external auditor, and consequences for the companies and individuals involved in the frauds.

Many of the characteristics noted by the Beasley et al. (1999) study could be of interest to researchers attempting to predict or detect fraud, but there are a few characteristics that are particularly interesting with respect to this study. First is the finding that some misstating companies were experiencing losses or were near break-even prior to the fraud. This indicates that a company's financial distress may be a precursor to a misstatement, supporting the use of the Altman Z-score as a predictor of misstatements. Second, Beasley et al. (1999) found that misstatements were most prevalent in the computer hardware and software, "other" manufacturing, and financial service industries, with these industries making up 35% of the financial statement frauds in the sample. Healthcare and health products, retailers and wholesalers, and "other" service providers comprised an additional 23% of the fraud companies in the sample. This may indicate that firms in these industries may be the most likely to misstate their financial statements, or it may indicate that the SEC focused their attention on these industries more than on others during the period 1987-1997. A limitation of this and other studies using AAERs as a data source is that this is an unknown: whether more fraud is perpetrated in any particular industry or whether the SEC places more of a focus on any particular industry. Third, the Beasley et al. (1999) study found that the misstatement of assets was the most frequent misstatement type within the sample, with revenue and net income misstatements following close behind.

The average misstatement of assets was nearly $40 million while the average misstatement of revenue and net income was nearly $10 million and about $16 million, respectively. Beasley et al. (1999) found that the asset accounts that were most frequently misstated were inventory, accounts receivable, and property, plant, and equipment. Fourth and finally, it is interesting to note the methods in which assets and revenue were frequently misstated. To overstate assets, the companies in the sample recorded fictitious assets or assets not actually owned by the company or capitalized items that should have been expensed. To overstate revenues, the companies in the sample recorded fictitious revenues or recorded revenue prematurely. Additional methods of misstatement were not available to the authors. Data detailing fraud methods or techniques may be useful in future research.

Following some very high-profile and very costly financial statement frauds that came to light immediately after the release of the initial COSO-sponsored report, COSO commissioned a second study of fraudulent financial statement occurrences, covering AAERs issued between January 1, 1998 and December 31, 2007. In this study, Beasley et al. (2010) updated the findings from the previous study and noted the differences and similarities between the two time periods.

The authors noted that several large company and large dollar frauds occurred in 2001 and 2002 (including Enron and WorldCom), which caused a drastic departure from the mean fraud size from the previous study. In this updated study, the review of AAERs found 347 companies that were cited for the violation of Rule 10(b)-5 of the 1934 Securities Exchange Act or Section 17(a) of the 1933 Securities Act during that time frame. Instead of taking a sample of these alleged fraud firms as was done in the prior study, the authors obtained detailed information regarding each of these 347 firms and their associated frauds, examining the same as well as additional characteristics pertaining to the company and management that were studied previously. The findings of this study included details about the occurrences of financial statement fraud, management’s tone at the top, the nature of the frauds, the role of the board of directors, related party transactions, auditor considerations, and consequences for individuals and firms engaged in fraud. Beasley et al. (2010) places a greater emphasis on corporate governance than was addressed in Beasley et al. (1999).

The findings in Beasley et al. (2010) are extensive, covering a multitude of topics that are interesting to fraud researchers, although a few items from the study are of particular interest to this study. First, the authors noted several motivating factors that were discussed by the SEC in the AAERs. Some of the most commonly cited reasons for committing financial statement fraud include meeting analyst expectations, meeting internal targets, concealing the company's deteriorating financial position, and meeting targets for management bonus payouts. Not all AAERs discussed the motivation behind the frauds, so the authors did not tally the frequency of each motivation. Some of these motivations could be studied empirically in the future, if the required data becomes available (e.g., meeting internal targets). Second, Beasley et al. (2010) found that, consistent with the previous study, misstatements were most prevalent in the computer hardware and software, "other" manufacturing, and healthcare and health products industries, with these industries making up 51% of the financial statement frauds in this new time period. Retailers and wholesalers, "other" service providers, and telecommunications comprised an additional 23% of the fraud companies. Third, the Beasley et al. (2010) study found that improper revenue recognition overtook the overstatement of assets as the most frequent misstatement type within the time period.

The mean misstatement of revenue was $455 million, up from a mere $10 million in the previous study. The mean misstatement of assets was $227 million, up from just $40 million in the earlier time period, indicating the potential importance of analyzing AAERs separately for the two time periods. These significant fluctuations are due mainly to the much larger size of the firms cited for financial statement fraud in the most recently studied time period. Beasley et al.

(2010) found that the asset accounts that were most frequently misstated continued to be inventory, accounts receivable, and property, plant, and equipment.

Fourth and finally, it is interesting to note the methods frequently used to misstate revenue and assets. To overstate revenues, the companies recorded fictitious revenues or recorded revenue prematurely. In the updated COSO report, Beasley et al. (2010) outlined the specific techniques used to overstate revenue, including sham sales, conditional sales, bill and hold transactions, improper cutoff of sales, and unauthorized shipments, among others. To overstate assets, the companies in this later time period continued the trend of recording fictitious assets or assets not actually owned by the company or capitalizing items that should have been expensed. No additional detail or techniques regarding the misstatement of assets were provided.

The Beasley et al. (1999) and Beasley et al. (2010) studies indicate that the most frequently misstated account is revenue, followed by assets, particularly inventory, accounts receivable, and property, plant, and equipment. In response to concerns that firms accelerated revenue to meet earnings targets, the SEC issued Staff Accounting Bulletin 101 (SAB 101) in 1999. Altamuro, Beatty, and Weber (2005) find that firms affected by SAB 101 used revenue to manage earnings in the period prior to adoption of SAB 101. Callen, Robb, and Segal (2008) find that firms that have losses are valued based on the level and growth of revenue, and are more likely to manipulate revenue in violation of GAAP. Marquardt and Wiedman (2004) find that firms issuing equity manage revenue upward. Bonner, Palmrose, and Young (1998) find that auditors are more likely to be sued in cases on enforcement actions involving fictitious revenue or premature revenue recognition.

The frequency with which the revenue account is used to perpetrate fraud/misstatements is so significant that SAS 99 (now AU-C 240) requires that auditors presume the existence of a fraud risk related to improper revenue recognition. While auditors may consider fraud risk in other accounts, revenue is the only account in which auditors are required to specifically consider the risk of fraudulent reporting. The characteristics noted by Beasley et al. (1999) and Beasley et al. (2010) may be developed into variables useful in the prediction of financial statement fraud (financial distress) or to further focus the attention of research (fraud industries, fraud accounts, and fraud methods).

The focus of this paper is to rely on financial and non-financial data that is readily available and to determine any significant trends or associations in the data that are useful in the prediction of misstating firms by incorporating industry, and useful in the prediction of misstated accounts.

3. HYPOTHESIS DEVELOPMENT

3.1 Prediction Models by Industry

Both the Beneish (1999) and Dechow et al. (2011) models were somewhat successful in identifying misstatement firms. Building upon these models, it is possible to use methodology similar to that used by Dechow et al. (2011) to extend the model to specific industries. This is important because the Beasley et al. (1999) and Beasley et al. (2010) studies indicate that fraudulent financial reporting is more likely to occur in certain industries: computer hardware and software, “other” manufacturing, financial service, healthcare and health products, retailers and wholesalers, “other” service providers, and telecommunications. These findings were corroborated in the Dechow et al. (2011) analysis indicating that the computer, durable manufacturing, retail, and services industries were the most frequently cited for misstatements in AAERs. Dechow et al. (2011) also noted that three industries in particular were over-represented, as compared to the percentage of industry makeup in the COMPUSTAT database. In order of over-representation, those industries are computer hardware and software, retail, and services. As such, I selected these three industries to analyze in order to determine the variables most significant in predicting financial statement misstatements, based on industry.

The industry in which a firm operates has been found to be an important factor not just in the frequency of misstatements, but also in the likelihood of litigation, whether against the firm or the auditor, following a misstatement (Bonner et al., 1998; Francis, Philbrick, and Schipper, 1994). Palmrose (1988) found that auditor litigation following a client misstatement occurs most frequently within the technology and financial services industries.

I expect that the industry distinction as developed in this research may result in finding some variables that have more predictive power for one industry than for another. In fact, it may be possible that a variable with no significant predictive power in one industry will have highly significant predictive power in another industry. This potential difference may be due to industry-specific characteristics. For instance, a variable that incorporates inventory may prove to be significant in the prediction of fraud for the retail industry, but may not be significant at all in the prediction of fraud for the service industry. By including or omitting certain variables from the prediction models, the prediction models restricted by industry may provide greater predictive ability than a more generalized model. The variable restrictions based on industry will allow the assessment of the effectiveness of the models in providing more accurate within-industry predictions, which motivates the first hypothesis:

> *H1: Prediction models individualized by industry will result in greater ability to predict misstating firms.*

3.2 Prediction Models by Account

Similar to the extension of industry-specific models, the previous prediction models can be extended to go beyond prediction of misstated firms to predict misstated accounts. Prior literature has shown that the most frequently misstated accounts have consistently been revenue, inventory, and accounts receivable (Beasley et al., 1999, Beasley et al., 2010). As such, I selected these three accounts to analyze in order to determine the variables most significant in predicting misstated accounts. In addition, AU-C 240 requires auditors to presume a fraud risk related to revenue (AICPA 2002). And finally, consistent with Beasley et al. (1999) and Beasley et al. (2010), Bonner et al. (1998) found that of the firms that were issued an AAER by the SEC, the misstated accounts that most frequently resulted in litigation were revenue and assets, although the individual misstated asset accounts that often led to litigation were not specified.

Based on the differing characteristics of each account, I expect that some variables will be significant in the prediction of fraud in one account while those same variables will have no significance in the prediction of fraud in other accounts. I expect that since the revenue and accounts receivable accounts are closely related, these two accounts may have similar findings: both accounts may include and exclude the same variables from the final prediction models. On the other hand, I expect that variables that incorporate inventory will be significant in the prediction of the inventory account being misstated, but perhaps not significant, or at least not to the same degree, for other misstated accounts. This motivates the second research hypothesis:

> *H2: Prediction models developed at the account level will help predict misstated accounts.*

3.3 Consistency of Predictive Variables

The Dechow et al. (2011) and Beneish (1999) models were developed using different variables and over different time periods. Very little is known about the stability of these models over time. Beasley et al. (1999) and Beasley et al. (2010) found similar trends across two different time periods. The same industries continued to be some of the most frequently cited for fraudulent financial statements and the same accounts continued to be some of the most frequently misstated. However, Kaminksi et al. (2004) did not find any ratios that were consistent predictors of fraud. These longitudinal results motivate the third research hypothesis:

> *H3: Predictive variables, whether specific to industry or account, will remain the same over time.*

Based on the Beasley et al. (1999) and Beasley et al. (2010) findings of similar misstatement-firm characteristics over time, I expect that the methodology for predicting misstated financial statements will remain stable across a longitudinal study of predictive variables. For example, I expect that if soft assets is a significant predictive variable in one time period, it will continue to be a significant predictive variable across most or all other studied time periods.

4. SAMPLE SELECTION

4.1 Data Source

Using data collected by Dechow (see full description of initial sample in Dechow et al., 2011), I use AAERs issued by the SEC to indicate firms that the SEC cited for financial statement misstatements. The AAERs document both intentional and unintentional GAAP violations by a public company and serve as the best proxy for misstatements. The instances of intentional deceit often result in the SEC's citation of Section 17(a) of the Securities Act of 1933 or Rule 10(b)-5 of the Securities Exchange Act of 1934, provisions enacted to specifically prohibit the inclusion or omission of information to fraudulently promote the purchase of a firm's stock by investors. Other potential data sources are available, including firms being sued by shareholders, firms with Sarbanes-Oxley Act internal control violations, and firms that have restated their financial statements. All of these data sources, including AAERs, are subject to the problem of selection bias. When it comes to selection bias in AAERs, I assume that the chance for Type I error (marking a non-misstating firm as a misstater) is very low, as the firms have been thoroughly investigated by the SEC. This is an advantage over a database of firms that have endured litigation initiated by the stockholders. For example, admission of guilt is rarely issued by accused firms, so it is difficult to assess which of these firms were misstating versus non-misstating. I conclude that the use of AAERs is advantageous over other available data sources, as has been done in previous research (Dechow et al., 2011; Beneish 1999; Kaminski 2004; Bayley and Taylor 2007; Brazel et al., 2009; Beasley et al., 1999; Beasley et al., 2010).

4.2 Timing of Misstatements

I start with the AAER data previously collected as described in Dechow et al. (2011). The AAER dataset includes all AAERs from the first AAER (AAER 1, released on May 17, 1982) through AAER 2261 (released on June 10, 2005). My initial sample was selected from this initial data. I created an expanded sample selected from additional AAER data collected by Dechow. This additional data includes AAER 2262 through AAER 3403 (released on August 31, 2012). The data is comprehensive, listing the cited firms, quarters and years of misstatement, misstatement accounts, etc.

Although I am unable to precisely reconcile the AAERs and firm-years that Dechow et al. (2011) used, the differences are minor. To provide more up-to-date analyses, I expand the time frame and consequently the number of AAERs that are included in the sample. A recent version of the Dechow data includes records through AAER 3403, which are releases dating through August 31, 2012. In this paper I run dual analyses: the first including only the AAERs included in the Dechow et al. (2011) study, and the second including the available AAER records through August 31, 2012. A second longitudinal study is also run, diving the misstatement-years into three decade-based time periods. These longitudinal analyses serve to provide evidence regarding the consistency of the model across time periods. In other words, the dual analyses show the consistency of the models across time periods to provide evidence on whether the characteristics of misstating firms have changed over time, and whether a prediction model can remain static over time or should be fluid, changing over time as the characteristics of misstating firms change.

Table 1 Panel A shows that there are 2,261 AAERs in the initial data set, however 74 AAERs are missing. An AAER may be missing for any number of reasons. Per Dechow et al. (2011), many of these AAERs were missing (19), intentionally omitted (11), did not involve specific company names (41), or were missing from the data set without a specified reason (3).

The expanded data set includes an additional 1,142 AAERs, of which 75 are missing for unspecified reasons, for a total of 3,254 AAERs in the expanded data set.

Panel B of Table 1 shows the distribution of AAERs by year. There were a large number of AAERs issued during years 2000-2004 and 2006-2009. These time frames coincide with the financial statement misstatements resulting in Sarbanes-Oxley regulations and the 2008 financial crisis, respectively. The number of AAERs for 2012 is significantly less than recent years because 2012 includes data for a partial year.

Table 1 Panel C shows that each firm may have more than one AAER, and conversely, each AAER may refer to more than one firm. For the initial and expanded samples, 41% and 43% of firms were issued just one AAER, respectively. One firm, Enron, was issued 24 AAERs by the cutoff date of the initial sample and 46 AAERs by the cutoff date of the expanded sample. Table 1 Panel C indicates 896 and 1,297 firms were cited by the SEC based on the 2,187 and 3,254 AAERs, by respective sample.

Table 2, Panels A and B show the number of misstating firms in each year for both the initial sample and the expanded sample, respectively. Panel A shows that the highest concentration of misstating firms centered around 1999 and 2000, the time-period immediately preceding the Sarbanes-Oxley legislation. Panel B shows that the highest concentration of misstating firms continues to be centered around 1999 and 2000, extending into 2001 and 2002 as well. In addition, the total number of misstating firms for 1999-2002 (as well as other years, to a lesser extent) increased with the expanded sample; adding in AAERs from 2005-2012 added a total of 633 misstating firms to the expanded sample compared to the initial sample. Since AAERs are issued after (sometimes years after) a misstatement occurs, this significant expansion of the initial sample is expected.

5. RESEARCH DESIGN AND EMPIRICAL RESULTS

5.1 Variables Analyzed

Closely following the Dechow model, I analyzed variables relating to accruals, financial performance, non-financial measures, and off-balance sheet information. Prior literature has shown that simple accrual calculations can have value in the prediction of misstatements. As a result, I include numerous variables calculated to measure the quality of accruals as well as changes in certain financial measures that can be manipulated through estimates (receivables, inventory, and soft assets, for example). Performance variables are used to determine if a firm may be motivated to misstate the financial statements because of worsening performance by manipulating other information in the financial statements (a leading indicator), or if these variables may themselves be manipulated as a way to hide worsening performance (a lagging indicator).

Non-financial measures are emerging as a new avenue to explore with regard to their predictive ability. A misstating firm may focus on manipulating the financial data to hide poor performance, but making corresponding manipulations of non-financial measures may be overlooked. For example, a firm may manipulate revenue, showing a large increase over the prior year, but fail to also manipulate the reported number of employees during the year, a non-financial measure that often corresponds to a firm's revenue. Finally, operating leases are often used by misstating firms to finance purchases of property, plant and equipment without showing a debt directly on the balance sheet. Therefore, the simple use of operating leases, or perhaps more importantly, the increased use of operating leases may be indicators of misstated financial statements. Appendix A offers a description of how each variable is calculated. All variables previously used in Dechow et al. (2011) are calculated in this paper using the same methods as in

Dechow et al. (2011). The data used in the calculation of each variable was obtained from COMPUSTAT.

5.2 New Variables Analyzed

As previously noted, the Dechow model and the Brazel model vary in their calculation of abnormal change in employees. This study tests each measure of the abnormal change in the number of employees to determine whether each is useful in the prediction of fraud, and which provides more predictive usefulness. The EMP DIFF variable used in Brazel et al. (2009) is calculated in this paper using the same method as in Brazel et al. (2009). See Appendix A for the calculation method for this and all variables.

For all industries, receivables as a percentage of total sales (rect_sales) is examined as a potentially significant variable in the prediction of fraud firms. Sales on account are more easily manipulated than cash sales, so I expect that as this ratio increases and/or increases as compared to other firms within the same industry, the likelihood of fraud will also increase. As total sales change year to year, we would expect the percent of those sales that are made on account to remain about the same. In addition, I would expect this ratio for any given firm-year to be comparable to the same ratio for other firm-years within the same industry. This paper examines anomalies in this ratio to determine if this variable is a reliable predictor of fraud.

For the computer and retail industries in particular, I examine the year-over-year change in total sales (ch_sales) and the change in receivables as a percentage of the change in total sales, expecting that as the percent change in total sales increases and as the percent change in receivables compared to the percent change in total sales increases, the likelihood of fraud also increases. By comparing these ratios for any particular firm-year to the same ratios for the industry as a whole, I can calculate abnormal changes in sales as compared to the industry. I expect that as abnormal changes in sales increase, so does the likelihood of fraud. Similarly, I expect that as abnormal changes in receivables to total sales increases, the likelihood of fraud does as well. These variables were selected for the computer hardware and software industry because many sales within this industry are business to business transactions, in which the purchases may be made on account, affecting the trade receivables account.

For the retail industry, several additional new variables based on sales are introduced. Some data specific to the retail industry can be obtained via COMPUSTAT: sales per square foot of retail space, sales per retail store, and same store sales (sss; this data item presents the sales figures only for the same store locations that existed in the prior year). From this data, I calculate change in sales (ch_sales), change in same store sales (ch_sss), change in net sales per square foot (ch_net_sales_sqft), and change in net sales per retail store (ch_net_sales_stores). I compare these variables for each firm-year to the corresponding variables for all firms within the industry, by year. I expect that large year-over-year and large abnormal changes as compared to the industry will be indicative of misstatement. These variables were selected specifically for the retail industry because the underlying data is specific to the retail industry; the retail industry is the only industry in which data on sales by store, by number of stores, or by retail square foot is collected and recorded. In addition, the number of stores and the number of retail square feet are considered nonfinancial measures. If a financial misstatement is made, comparison to a related nonfinancial measure may result in anomalies that could signal a potential misstatement.

It is important to note that the data for these variables was very limited in COMPUSTAT; very few retail firm-years, relatively, have this data. As a result, the N for these regressions is very small, relative to the entire data set. In order to run the required regressions, the missing data were imputed statistically. A second item to note regarding the sales data for the retail industry is that the sales data presented in sales per square foot, sales per store, and same store sales does not agree to the sales data that is used to calculate other variables in this paper. While the data used to calculate the variables differs between these retail variables and all other general variables in which a sales figure is used, the reliability of the data is not impacted. The new retail variables all use the same sales figure and are therefore comparable.

The service industry is greatly varied with respect to the types of firms included. As a result, no new variables beyond rect_sales and rect_sales_indservices were added to the model.

5.3 Analysis of Misstating Firm-Years

I compare the misstated firm-years to all firm-years listed in COMPUSTAT between 1971 and 2003 (initial sample), and between 1971 and 2011 (expanded sample). While there were some AAERs issued for early and recent firm-years, I excluded firm-years with fewer than ten in any particular year. This means that for the initial sample, firm-years 1971 through 1979 and 2003 were excluded. For the expanded sample, firm-years 1971 through 1979 and 2009 through 2011 were excluded. COMPUSTAT data contained about 213,000 firm-years between 1980 and 2008. I reduced the number of firm-years in the data first by removing the firm-years with no total assets reported (none reported, as opposed to $0 reported), and then by removing the firm-years associated with banks and insurance (SIC codes 6000-6999). Any remaining duplicate firm-years were also removed from the COMPUSTAT data. Finally, outliers were removed from the data after individual review. The remaining data consist of 145,199 (181,696) total firm-years, of which 834 (1068) are misstatement firm-years for the initial (expanded) samples. For the computer hardware and software industry, 19,213 (25,491) firm-years are included in the data, of which 244 (328) are initial (expanded) misstatement firm-years. For the retail industry, 10,725 (12,907) firm years are included in the data, of which 59 (76) are initial (expanded) misstatement firm years. Finally, for the services industry, 17,981 (22,412) firm-years are included in the data, of which 109 (149) are initial (expanded) misstatement firm years.

In order to determine the variables significant in predicting misstatement by industry, I used a dummy with a value of 1 indicating a misstatement in any given firm-year, 0 if no misstatement. To indicate misstated accounts, I used a dummy with a value of 1 to indicate misstatement in each of the revenue, accounts receivable, and inventory accounts, 0 if no misstatement in each of the three accounts. The logistic model used in each of the base regressions (prior to adding in the new variables) is:

Misstating Firm-Year = wc_acc + rsst_acc + ch_rec + ch_inv + soft_assets + ch_cs + ch_cm + ch_roa + ch_fcf + tax + ch_emp (EMP DIFF) + leasedum + oplease + issue + cff + leverage + bm + ep

5.4 Prediction Models by Industry

In order to determine the variables significant in predicting misstatement when restricting by industry, I selected three frequently misstating industries to analyze: computer hardware and software, retail, and services. Using only the firm-years within each of the three industries selected, I ran separate logistic regressions for each industry. Tables 3 and 4 show the results by initial and expanded sample and by industry. The first set of results in each table reports the results of the Dechow model applied to all industries for each time period and is the same across each panel. The second set reports the results of the Dechow model applied to observations in the specific industry. The third model is the Dechow model plus added variables applied to the industry. Table 3 reports results for the initial sample period of 1980-2002, and Table 4 reports the results for the full sample period 1980-2008.

The logistic analysis of the computer hardware and software industry found that for both the initial and expanded samples, ch_rec, soft_assets, leasedum, oplease, issue, and leverage were significant in predicting a misstated firm-year (Tables 3 and 4, Panel A). These results differ from those in the Dechow model run on all industries; while Dechow et al. (2011) found rsst_acc, ch_inv, ch_cs, ch_roa, and ch_emp significant, these variables are not significant for the computer hardware and software industry. On the other hand, leasedum and leverage are significant for the computer hardware and software industry, but were not found to be significant in the Dechow model run on all industries. None of the new variables analyzed were found to be significant, indicating that none of the new variables for the computer hardware and software industry are useful predictors of misstatements. While no *new* variables were significant for this industry, it is interesting to note that there are some differences in variables that are useful for fraud prediction in general, as compared to fraud prediction in this industry alone, especially the importance of leverage in the prediction of fraud in the computer hardware and sales industry.

Tables 3 and 4 Panel B shows the regression results for the retail industry. I removed firm-years with SIC codes between 5000 and 5190 from the industry analysis, as those codes are for various wholesalers, as opposed to retail operations that involve selling directly to customers. The logistic analysis of the retail industry found that for both the initial and expanded samples, only soft assets and issue were significant in predicting a misstated firm-year, prior to adding the new variables to the regression. After adding the new variables to the regression, ch_sss and ch_rec were found to be significant. These results differ from those in the Dechow model run on all industries; while Dechow et al. (2011) found rsst_acc, ch_inv, ch_cs, ch_roa, ch_emp, and oplease significant, these variables were not significant in a regression specific to the retail industry. In addition, ch_rec was not a significant predictor of misstatement for retail firms until new variables were added to the regression. In that case, ch_rec was significant, as was ch_sss, a new variable. Interestingly, change in receivables was positively associated with misstating firms, suggesting the existence of misstated revenue. However, the change in same store sales was negatively associated with misstating firms, possibility indicating financial distress. No other new variables were found to be significant for the retail industry. However, the explanatory power of the model with added variables was higher than the Dechow model alone.

The logistic analysis of the service industry (Tables 3 and 4 Panel C) found that for both the initial and expanded samples, ch_inv, soft assets, ch_cs, leasedum, and issue were the significant variables in predicting a misstated firm-year. As with the other industries, these results differ from those in the Dechow model run on all industries. Dechow et al. (2011) found rsst_acc, ch_rec, ch_roa, and ch_emp to be significant, although those variables are not significant when restricted to firms in the service industry. Neither of the two new variables were significant for the service industry. It is interesting to note that the power of the models increases when restricted to service firms, as opposed to all firms. R-squared increases from 0.042 (0.040) to 0.051 (0.055) when adding the restriction to service firms only in the initial (expanded) sample. This may mean that misstated firm-years within the service industry could be predicted more accurately than firms in other industries.

I ran each model twice: once using the variable ch_emp, found significant in Dechow et al. (2011), and once using the variable EMP_DIFF, found significant in Brazel et al. (2009). Running these two variables separately eliminated any collinearity that would have been present had both variables been run in the models simultaneously. In all industry and account models and in both the initial, expanded, and decade-based time periods in this study, neither ch_emp nor EMP_DIFF were found to be significant predictors of misstating firms, though EMP_DIFF usually outperformed ch_emp. This finding is unexpected, suggesting that changes in employees is not a significant predictor of misstatements. Further research will be required to reconcile the differences between this study and prior literature. It may also be interesting to determine if other non-financial variables found significant in Brazel et al. (2009) will be found to be useful predictors of misstatement when included in models such as those employed in this study.

After finding the significant variables for each industry and sample, I ran final regressions inclusive of only the significant variables. Without exception, the R-squared decreased from that of the third models in each panel, although for the services industry, the R-squared remained above that of the original Dechow model. This result is promising; for the service industry, at least, the significant variables found in this study offer better predictive ability than the Dechow model.

I note that the variables soft_assets and issue were highly significant across all three industries. This may indicate that these variables could be the most useful predictors of misstatement, regardless of industry. Soft_assets is a measure of assets that could be manipulated through management estimates (accounts receivable and the allowance for doubtful accounts, for example), as a percentage of total assets. Because of the opportunity to manipulate soft assets, it is expected that such a variable could be a consistent indicator of misstatement. Issue is a dummy variable indicating whether or not the firm issued securities during the firm-year. This is essentially a measure of a firm's need for capital, which could be raised alternatively through debt. This measure could indicate the firm's expansion or the firm's distress. In either event, raising capital through stock issuance is shown to be a reliable indicator of misstatement across all three industries.

Beasley et al. (1999) and Beasley et al. (2010) compiled misstatement information for the years 1987 to 1997 and 1998 to 2007, respectively, finding in part that computer hardware and software, retail, and service industries were among the most frequently cited for fraudulent financial statement reporting. Dechow et al. (2011) include similar findings as part of that study. Models to predict misstatements by firms within these specific industries could improve information for financial statement users, and this study finds evidence indicating that some industry-specific models outperform a more generalized prediction model. Noting the small number of misstating firm-years in the samples relative to the number of non-misstating firm-years in the samples, the lack of meaningful change in the R-squared for the models may be due to lack of power. For example, the sample representing the computer hardware and software industry includes misstating firm-years that make up just 1.3 percent of the total sample. For the retail and service industries, the misstating firm-years make up less than one percent of the total sample. Future research may consider using a matched sample approach to determine the most meaningful variables for use in misstatement prediction models.

I analyzed variance inflation factors (VIF) to determine whether the results were impacted by multicollinearity due to the added variables. VIFs exceeding 10, indicating potentially harmful collinearity, were present only in the retail model, in which rect_sales and rect_sales_indretail were highly correlated. To address this issue, I ran the retail regressions including only one of these two variables at one time. The panels for the retail industry model show the variable coefficients and p-values using rect_sales in the model.

5.5 Prediction Models by Account

In order to determine the variables significant in predicting misstatement by account, I used a dummy to indicate a misstatement in any given account, noting that a misstatement firm-year can have misstatements in multiple accounts. I ran separate logistic regressions for each account following the same base model used for the industry regressions. While additional accounts were cited by the SEC, I focus only on revenue, inventory, and accounts receivable, which are the most frequently misstated accounts. Table 5 Panels A through C show the results of the regressions by misstated account.

The logistic analysis of the revenue account found that across all time periods, ch_rec, soft_assets, and issue were significant (Table 5 Panel A). Similar to the findings in the industry analyses, the secondary period for the revenue account failed to show significance for ch_inv, leasedum, and oplease, even though these three variables were significant for the initial and "all years" time periods. The R-squared values for the revenue account analysis are higher than the base model run across all years and all accounts, indicating that a revenue-only analysis may be a more powerful predictor of misstatement than the generalized model.

The logistic analysis of the accounts receivable account found that across all time periods, ch_rec and soft_assets were significant (Table 5 Panel B). With the exception of the secondary time period, ch_inv, leasedum, oplease, and issue were significant. These results are very similar to the results for the revenue account, including the higher R-squared for all time periods except for the secondary period, when compared to the analysis across all years and all misstatement accounts.

Finally, the logistic analysis of the inventory account found that ch_inv and soft_assets were significant in predicting a misstatement across all time periods (Table 5 Panel C).

Leasedum and issue were significant in all time periods except for the secondary time period. Interestingly, ch_rec and oplease were found to be significant in the revenue and accounts receivable analyses but not in the inventory analysis. Since the inventory models outperform the generalized model across all years and all accounts, this may indicate that misstatement in the inventory account can be more accurately predicted than misstatements in the generalized model.

I analyzed variance inflation factors (VIF) to determine if the results were impacted by multicollinearity due to the added variables. For the account-specific models, there were no VIFs exceeding 10, indicating the absence of potentially harmful collinearity.

5.6 Longitudinal Study of Predictive Variables

I initially selected three time periods over which to analyze the consistency of the variables found to be significant predictors of financial statement misstatements. The initial time period of 1980 through 2002 coincides with the time period used in Dechow et al. (2011). The secondary period of 2003 through 2008 extends from the end of the initial time period through the end of the available AAER data. The time period of 1980 through 2008 encompasses both the initial and secondary time periods to get an overall picture of the significant predictive variables. Table 6 shows the results of the longitudinal study of predictive variables, for all industries as well as for and the three industries analyzed, individually.

A logistic regression model inclusive of all industries (Table 6 Panel A) found that ch_rec, soft_assets, leasedum, and issue were significant across all three time periods, often maintaining the same level of significance and similar coefficients across the time periods, as well. The analysis of the secondary time period found that while ch_inv and ch_cm were significant in the initial and “all years” periods, they were not significant in the secondary period.

The opposite is true of cff and leverage; these two variables were significant in the secondary time period but not in the initial or “all years” periods.

For the computer hardware and software industry (Table 6 Panel B), only the variable “issue” was a significant predictor across all time periods. The secondary time period found ch_rec, soft_assets, leasedum, oplease, and leverage to be not significant, even though these variables were significant in the initial and “all years” time periods. The variables cff and rect_sales_indcomp were significant predictors in the secondary period even though they were not significant in the other two time periods.

Table 6 Panel C shows that for the retail industry, only soft_assets was significant across all three time periods. Ch_rec, issue, and ch_sss were significant across the initial and “all years” periods but were not significant for the secondary period. There were no variables significant in the secondary period but not in the remaining time periods.

Finally, Table 6 Panel D shows that for the services industry soft_assets and issue were significant across all three time periods, while ch_inv, ch_cs, and leasedum were significant across all time periods except for the secondary time period. There were no variables significant in the secondary period but not in the remaining time periods. The variables soft_assets and issue were most consistently significant across the time periods and all industries, indicating that these two variables in particular may be the most reliable predictors of misstatements for any given industry and any given time period.

The findings of this first longitudinal study showed consistency between the initial and “all year” periods with regard to the variables that were found to be significant predictors of misstatement. The differences across time periods lay solely with the secondary time period, whether the secondary time period showed some variables only significant in that time period or significant in every other time period except for the secondary period. Some of the differences in significant variables across time periods may be a result of the difference in the amount of time included in the time periods; the initial time period consists of 23 years while the secondary time period consists of only 6 years. The disparity between the time periods may not allow for enough data in the secondary time period to be comparable to the initial time period.

To allow for more a more consistent comparison of the variables over time, I conducted the second portion of the longitudinal study using time periods based on natural decades. Using the periods 1980-1989, 1990-1999, and 2000-2008, I am able to more reliably determine the variables that are consistently significant across time. The third decade analyzed is short by one year as that is the final year of data available. These time periods presented a statistical challenge at times, however, with very few misstatement firm-years included an any given decade, especially when also restricted by industry or account. There were some regressions that would not run when using a logistic regression, but I found that using a backward stepwise logistic regression helped to alleviate the problems encountered when using a logistic regression. Therefore, I used the backward stepwise logistic regression for all models in this second portion of the longitudinal study.

Table 7 and Table 8 show the longitudinal results of regressions by industry and by account, with time periods arranged by decade instead of by initial and expanded samples. Table 7 Panel A and Table 8 Panel A show the results of all industries and all accounts, by decade. The findings show that ch_rec, ch_inv, soft_assets, leasedum, and issue were the most consistent predictors of misstatements across all three time periods, without regard for industry or misstated account.

Table 7 Panel B shows that for the computer hardware and software industry, ch_rec, soft_assets, and leasedum were the most consistent predictors of misstatement across time. As this industry engages in a significant amount of business to business sales, it is expected that trade receivables may make up a large portion of the sales, so it makes sense that abnormal changes in trade receivables is a consistent predictor of misstatements for this industry. It follows that the ratio of soft assets to total assets is also a consistent predictor of misstatements for the computer hardware and software industry, since trade receivables are included as a soft asset. Other soft assets may include inventory, supplies, prepaids, and goodwill, many of which can be manipulated and still pass a reasonability test. The existence of an operating lease obligation is the third and final consistent predictor of misstatements for the computer hardware and software industry, and like the first two predictors, is also related to assets. An operating lease will not result in the associated asset being included on the balance sheet like a capital lease would. If a firm within this industry has an operating lease obligation, the likelihood of misstatement increases.

For the retail industry (Table 7 Panel C), there were no consistent predictors of misstatement across time. This may be due in part to the small number of misstating firm-years within the retail industry, with only 9 in the first period, 31 in the second period, and 36 in the third period, for a total of just 76 misstating firm years across three decades. Another factor to consider is the changing "storefront" of the retail industry. With online sales making up an increasing percentage of total retail sales as compared to sales made in brick-and-mortar stores, measures such as same store sales, sales per number of stores, and sales per square foot of retail space will likely become irrelevant within the retail industry. In the future, perhaps measures such as sales per transaction or sales per website hit may replace the obsolete measures.

For the services industry (Table 7 Panel D), soft_assets and issue were the most consistent predictors of misstatements. I note that the change in receivables variable is not a consistent predictor of misstatement for this industry even though it is consistent across all time periods in the general model. This makes sense as sales transactions on account are less frequent in the services industry: often payment is due before or at the time that services are rendered. The soft assets ratio is a consistent predictor, however. It is likely that assets other than cash, receivables, and PP&E account for the significance of this variable. Finally, the issuance of securities during the year is a consistent predictor of misstatement for the services industry. The issuance of securities may indicate a cash shortage or the need for funding to allow for growth. If a company is looking to grow, there may be incentive to manipulate or misstate the financial statements to make the company appear to be in a better position than it actually is.

For the revenue account, Table 8 Panel B shows that ch_rec, ch_inv, soft_assets, and issue were the most reliable predictors of misstatements. Similar results were found for the misstatements in the accounts receivable account (Table 8 Panel C), with the exception that the issue variable was not found to be consistently significant for accounts receivable. It was expected that the same variables that were significant predictors of the revenue account would also be significant predictors for the accounts receivable account, as the two accounts are so closely linked: revenue transactions are frequently completed on account, and in the event that revenue is intentionally overstated by creating false sales transactions, the accounts receivable account is usually overstated as well. It follows that change in receivables would be a consistent predictor of misstatements in both the revenue and accounts receivable account, and as receivables are a soft asset, the ratio of soft assets to total assets being a consistent predictor of misstatements makes sense as well.

Table 8 Panel D addresses predictors of misstatements in inventory. Following the rationale from the prior paragraph for revenue and accounts receivable, changes in inventory and soft assets were consistent predictors of misstatement.

Anecdotally, anomalies in inventory often coincide with anomalies in receivables and in the soft assets ratio. For example, if a false sale of a physical good is recorded, revenue will be overstated, accounts receivable will likely be overstated, and if inventory is not manipulated as well, the inventory account will likely be abnormally low or high, compared to the prior year in which no misstatements were made. In addition, intentional misstatements in the inventory account often must be hidden by maintaining or increasing the total amount of the misstatement in each subsequent year. As a result of revenue misstatements touching so many different accounts, it makes sense to see that abnormal changes in inventory and an abnormal soft assets ratio are consistent predictors of misstatement for the revenue, accounts receivable, and inventory accounts.

Compared to the general model of all industries, only the longitudinal model for the services industry consistently matched or outperformed the general model, as evidenced by the R-squared values (Table 7 Panel D). In addition, the revenue model and the inventory model consistently matched or outperformed the general model of all misstated accounts (Table 8 Panel B and Table 8 Panel D, respectively).

The results of the analyses of time periods by decade coincide with the results of the analyses of the time periods of "initial" versus "expanded," corroborating the finding that many predictive variables are consistent across time periods, regardless of how the time periods are delineated.

6. CONCLUSIONS AND DISCUSSION

In this study, I used Accounting and Auditing Enforcement Release (AAER) data compiled by Dechow et al. (2011) to test variables for their ability to predict misstating firm-years within the computer hardware and software, retail, and service industries. I separated the firm-years into the initial and expanded samples based on the time period used in Dechow et al. (2011) and recently available data. Using 181,696 firm-years, I tested variables previously included in a general model, adding an alternative variable for abnormal changes in the number of employees.

I then analyzed the firm-years within the computer hardware and software, retail, and service industries. For each of these industries, I tested the variables from the general model, and then added new, previously untested variables to the industry models.

Based on these general and industry-specific models, I found that two measures of abnormal changes in the number of employees were not significant predictors of misstatement, even though previous research did find them to be significant in general prediction models. I also found that for the retail industry, the added variable “change in same store sales” was a significant predictor of misstatement, and this industry-specific model had more explanatory power than the general model. The industry-specific models for the computer hardware and software industry and the service industry did not identify additional new significant variables, and did not have greater explanatory power than the general model.

I also found that some variables that Dechow et al. (2011) found to be significant predictors of misstatements in a general model were not significant in the industry-specific models tested. In contrast to this, I also found that some variables that were not previously found to be significant predictors in a general model were significant predictors in some industry-specific models.

The greater explanatory power of the retail-specific model may be helpful to auditors in the planning stage of the audit, during which time the risk of misstatement is assessed and responses to that risk are determined. For audit clients within the retail industry, this model could assist the auditor in determining if a client presents a greater risk of misstatement. Similarly, government investigators could use this industry-specific model as a method of determining high-risk firm-years and select firm-years for review based in part on this assessment. Of course, creating industry-specific models with high explanatory power for more industries will be ideal and may be an avenue for future research.

This study also found that across all industries and across all time periods, soft_assets and issue are the most reliably significant and likely the most useful predictors of misstatement. Several other variables are significant across the initial and “all years” time periods. The secondary time period is much shorter than the initial time period, which may be a factor in the inconsistency of the significance of these other variables.

Further, I analyzed the variables in the base model to determine which, if any, are significant in the prediction of misstated accounts, looking specifically at the revenue, accounts receivable, and inventory accounts. I tested the models across the initial, secondary, and “all years” time periods to determine if the significant predictor variables changed over time. I found that in all three models, ch_inv, soft_assets, leasedum, and issue were significant, although the significance often lacked in the secondary time period. Ch_rec and oplease were significant in the revenue and accounts receivable models, but were not significant in the inventory model. All three models outperformed the generalized model, indicating that the account-specific models may predict a misstated account with more precision and power than the generalized misstatement model. I also analyzed the models across decade-based time periods, finding similar results to those found in the first portion of the longitudinal study.

This study is limited first by the data. In investigating and issuing AAERs, the SEC does not catch all instances of misstatement. Some misstatements may be small enough to fly under the radar and not be caught by the SEC, and at the same time, the SEC cannot investigate every firm that is determined to be of higher risk. This selection bias by the SEC may have a significant impact on the accuracy of the prediction models, with the potential for many false positives: firms determined to be high-risk but not found to have misstated financial statements.

This study is also limited by the small number of misstating firm-years, as compared to the relatively large number of non-misstating firm-years. This disparity may result in low power for each of the models. Future research may benefit from running similar models to determine significant predictors of misstatement, using a matched sample approach. This approach may eliminate the power issue and more precisely determine the indicators of financial statement misstatements.

Finally, future research may include the further study of the significant predictive variables over time: are the predictors of misstatements static, or are they fluid and change as fraud methods and industry trends change? Do the coefficients change over additional time periods? Using similar statistical analyses to predict the method used to misstate the financial statements (channel stuffing or bill-and-hold, for example) may also be an interesting area for future research.

Table 1
Sample Description

Panel A: Sample selection of Accounting and Auditing Enforcement Releases (AAERs)

Number of AAERs		
AAER No. 1 - No. 2261 from May 1982 to June 2005	2,261	
Less: missing AAERs	(74)	
Total AAERs for initial sample period	2,187	
AAER No. 2262 - No. 3403 from June 2005 through August 2012		1,142
Less: missing AAERs		(75)
Total AAERs for added period		1,067
Total sample of AAERs		3,254

Panel B: Frequency of AAERs by year

AAER release date	Number of AAERs	Percentage	AAER release date	Number of AAERs	Percentage
1982	2	0.1	1998	85	2.6
1983	16	0.5	1999	111	3.4
1984	28	0.9	2000	142	4.4
1985	35	1.1	2001	125	3.8
1986	39	1.2	2002	209	6.4
1987	51	1.6	2003	237	7.3
1988	37	1.1	2004	209	6.4
1989	38	1.2	2005	187	5.8
1990	35	1.1	2006	171	5.3
1991	61	1.9	2007	214	6.6
1992	78	2.4	2008	146	4.5
1993	76	2.3	2009	172	5.3
1994	120	3.7	2010	117	3.6
1995	107	3.3	2011	111	3.4
1996	121	3.7	2012	43	1.3
1997	134	4.1			
			Total	**3254**	**100.0**

Panel C: Frequency of AAERs by firm - initial and expanded samples

Number of AAERs for each firm	Initial Sample		Expanded Sample	
	Number of firms	Percent of firms	Number of firms	Percent of firms
1	371	41.4	561	43.3
2	234	26.1	310	23.9
3	108	12.1	155	12.0
4	70	7.8	102	7.9
5	40	4.5	62	4.8
6	32	3.6	30	2.3
7	14	1.6	20	1.5
8	10	1.1	21	1.6
9	3	0.3	7	0.5
10	6	0.7	9	0.7
11	2	0.2	4	0.3
12	2	0.2	1	0.1
13	1	0.1	5	0.4
14	0	0.0	3	0.2
15	0	0.0	1	0.1
16	0	0.0	1	0.1
17	0	0.0	1	0.1
18	0	0.0	1	0.1
20	1	0.1	1	0.1
15	1	0.1	0	0.0
24	1	0.1	0	0.0
25	0	0.0	1	0.1
46	0	0.0	1	0.1
Total	**896**	**100.0**	**1297**	**100.0**

Table 2
Frequency of Misstating Firms by Calendar Year

Panel A: Distribution of misstated firm years - initial sample

Year	Firm-years	Percentage	Year	Firm-years	Percentage
1971	1	0.11	1987	24	2.75
1972	1	0.11	1988	27	3.09
1973	1	0.11	1989	43	4.92
1974	2	0.23	1990	34	3.89
1975	2	0.23	1991	45	5.15
1976	1	0.11	1992	48	5.49
1977	1	0.11	1993	42	4.81
1978	4	0.46	1994	35	4.00
1979	9	1.03	1995	37	4.23
1980	17	1.95	1996	39	4.46
1981	23	2.63	1997	45	5.15
1982	31	3.55	1998	57	6.52
1983	25	2.86	1999	72	8.24
1984	25	2.86	2000	68	7.78
1985	17	1.95	2001	39	4.46
1986	30	3.43	2002	21	2.40
			2003	8	0.92
			Total	**874**	**100.00**

Panel B: Distribution of misstated firm years - expanded sample

Year	Firm-years	Percentage	Year	Firm-years	Percentage
1971	1	0.07	1991	46	3.05
1972	1	0.07	1992	49	3.25
1973	1	0.07	1993	44	2.92
1974	2	0.13	1994	39	2.59
1975	2	0.13	1995	42	2.79
1976	1	0.07	1996	46	3.05
1977	1	0.07	1997	64	4.25
1978	4	0.27	1998	83	5.51
1979	9	0.60	1999	111	7.37
1980	17	1.13	2000	130	8.63
1981	24	1.59	2001	123	8.16
1982	32	2.12	2002	106	7.03
1983	26	1.73	2003	92	6.10
1984	26	1.73	2004	71	4.71
1985	18	1.19	2005	56	3.72
1986	31	2.06	2006	33	2.19
1987	26	1.73	2007	28	1.86
1988	27	1.79	2008	11	0.73
1989	43	2.85	2009	4	0.27
1990	34	2.26	2010	2	0.13
			2011	1	0.07
			Total	**1507**	**100.00**

Table 3
Regression Results by Industry: 1980 – 2002

Panel A: Computer Hardware and Software									
	Dechow Model - All Ind.			Dechow Model - CH&S			Dechow Model PLUS - CH&S		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.8641	0.0000	***	-7.4841	0.0000	***	-7.4785	0.0000	***
wc_acc	-0.0182	0.3755		-0.0767	0.4646		-0.0735	0.5034	
rsst_acc	0.0176	0.1873		0.0548	0.1494		0.0521	0.1787	
ch_rec	1.2916	0.0000	***	1.3065	0.0024	**	1.5880	0.0012	**
ch_inv	1.0549	0.0009	***	0.8364	0.2077		0.8963	0.1891	
soft_assets	1.7432	0.0000	***	1.3629	0.0000	***	1.3450	0.0000	***
ch_cs	-0.0008	0.0368	*	0.0000	0.9887		0.0008	0.8094	
ch_cm	-0.0007	0.0004	***	0.0004	0.9203		0.0004	0.9165	
ch_roa	-0.0027	0.4335		-0.0121	0.6038		-0.0105	0.6942	
ch_fcf	0.0000	0.8457		0.0000	0.9904		0.0000	0.9867	
tax	-0.1274	0.6460		-0.2485	0.2787		-0.2667	0.2477	
ch_emp	0.0004	0.8354		0.0017	0.8435		0.0015	0.8685	
EMP DIFF	-0.0034	0.4805		-0.0043	0.5809		-0.0014	0.9170	
leasedum	0.6373	0.0000	***	0.7398	0.0108	*	0.7456	0.0107	*
oplease	-0.2640	0.1217		-0.6698	0.0019	**	-0.6667	0.0020	**
issue	1.2533	0.0000	***	1.6566	0.0001	***	1.6470	0.0001	***
cff	0.0094	0.7486		-0.1499	0.2625		-0.1342	0.3219	
leverage	-0.0020	0.8879		0.1136	0.0008	***	0.1136	0.0008	***
bm	-0.0008	0.6178		-0.0002	0.9406		-0.0002	0.9407	
ep	0.0007	0.4115		0.0206	0.1318		0.0202	0.1447	
rect_sales							0.0423	0.9235	
rect_sales_indcomp							-0.0048	0.9710	

ch_sales			-0.0305	0.3832
ch_sales_indcomp			-0.0025	0.9322
ch_rect_ch_sales			-0.0341	0.2674
ch_rect_ch_sales_indcomp			0.0006	0.6401
Misstating firm-years	834	244		244
Nonmisstating firm-years	144,365	18,969		18,969
	145,199	19,213		19,213
R-Squared	0.042	0.038		0.039

Panel B: Retail

	Dechow Model - All Ind.			Dechow Model - Retail			Dechow Model PLUS - Retail		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.8641	0.0000	***	-6.6803	0.0000	***	-6.5899	0.3173	
wc_acc	-0.0182	0.3755		-0.4847	0.5807		-0.6117	0.4682	
rsst_acc	0.0176	0.1873		0.3566	0.4412		0.4230	0.3428	
ch_rec	1.2916	0.0000	***	2.3056	0.0846		3.4247	0.0339	*
ch_inv	1.0549	0.0009	***	0.5348	0.6777		0.2973	0.8172	
soft_assets	1.7432	0.0000	***	1.3204	0.0369	*	1.6847	0.0137	*
ch_cs	-0.0008	0.0368	*	0.0007	0.9132		0.0006	0.9283	
ch_cm	-0.0007	0.0004	***	0.0002	0.9598		0.0010	0.9384	
ch_roa	-0.0027	0.4335		-0.6807	0.0917		-0.7147	0.0617	
ch_fcf	0.0000	0.8457		0.0217	0.1485		0.0228	0.1405	
tax	-0.1274	0.6460		-0.3028	0.6934		-0.2928	0.7232	
ch_emp	0.0004	0.8354		0.0176	0.0969		0.0162	0.2258	
EMP DIFF	-0.0034	0.4805		-0.0238	0.1120		-0.0198	0.2654	
leasedum	0.6373	0.0000	***	-0.4471	0.4084		-0.5450	0.3211	
oplease	-0.2640	0.1217		0.4829	0.5446		0.3113	0.6720	
issue	1.2533	0.0000	***	1.2434	0.0444	*	1.2595	0.0445	*
cff	0.0094	0.7486		0.3395	0.3604		0.3631	0.3242	
leverage	-0.0020	0.8879		-0.0155	0.8907		-0.0029	0.9800	
bm	-0.0008	0.6178		-0.0141	0.4738		-0.0118	0.5465	
ep	0.0007	0.4115		0.0081	0.4878		0.0104	0.2727	
rect_sales							-10.3373	0.0577	
rect_sales_indretail							0.5808	0.0762	
ch_sales							-0.0324	0.6916	
ch_sales_indretail							0.0252	0.4404	
sss							-0.0886	0.3670	

sss_indretail							0.0951	0.3623	
ch_sss							-0.2719	0.0007	***
ch_sss_indretail							-0.0651	0.2260	
ch_net_sales_sqft							1.6252	0.9382	
ch_net_sales_sqft_indretail							-0.4148	0.2885	
net_sales_stores							0.0000	0.9325	
net_sales_stores_indretail							-0.3200	0.9580	
ch_net_sales_stores							8.6217	0.4096	
ch_net_sales_stores_indretail							0.1595	0.4421	
Misstating firm-years		834			59			59	
Nonmisstating firm-years		144,365			10,666			10,666	
		145,199			10,725			10,725	
R-Squared		0.042			0.034			0.057	

Panel C: Services									
	Dechow Model - All Ind.			Dechow Model - Services			Dechow Model PLUS - Services		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.8641	0.0000	***	-8.0140	0.0000	***	-7.9681	0.0000	***
wc_acc	-0.0182	0.3755		-0.0104	0.8805		-0.0114	0.8720	
rsst_acc	0.0176	0.1873		0.0148	0.6738		0.0158	0.6437	
ch_rec	1.2916	0.0000	***	0.8342	0.1514		0.8196	0.1600	
ch_inv	1.0549	0.0009	***	2.0068	0.0083	**	2.0093	0.0082	**
soft_assets	1.7432	0.0000	***	1.7099	0.0000	***	1.7111	0.0000	***
ch_cs	-0.0008	0.0368	*	-0.0009	0.0185	*	-0.0009	0.0181	*
ch_cm	-0.0007	0.0004	***	-0.0003	0.6920		-0.0003	0.6863	
ch_roa	-0.0027	0.4335		-0.0054	0.8198		-0.0056	0.8095	
ch_fcf	0.0000	0.8457		0.0000	0.8620		0.0000	0.8548	
tax	-0.1274	0.6460		0.5835	0.1671		0.5836	0.1664	
ch_emp	0.0004	0.8354		-0.0004	0.8839		-0.0004	0.8864	
EMP DIFF	-0.0034	0.4805		-0.0032	0.7256		-0.0032	0.7264	
leasedum	0.6373	0.0000	***	0.7988	0.0263	*	0.7964	0.0277	*
oplease	-0.2640	0.1217		0.0444	0.9038		0.0434	0.9062	
issue	1.2533	0.0000	***	1.2761	0.0028	**	1.2706	0.0029	**
cff	0.0094	0.7486		0.0185	0.7554		0.0182	0.7612	
leverage	-0.0020	0.8879		0.0308	0.2583		0.0308	0.2583	
bm	-0.0008	0.6178		-0.0008	0.7761		-0.0008	0.7765	
ep	0.0007	0.4115		0.0037	0.5616		0.0037	0.5617	
rect_sales							-0.0864	0.6936	
rect_sales_indservices							0.0366	0.6569	
Misstating firm-years		834			109			109	
Nonmisstating firm-years		144,365			17,872			17,872	

	145,199	17,981	17,981
R-Squared	0.042	0.051	0.051

Table 4
Regression Results by Industry: 1980 – 2008

Panel A: Computer Hardware and Software									
	Dechow Model - All Ind.			Dechow Model - CH&S			Dechow Model PLUS - CH&S		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.8690	0.0000	***	-7.1813	0.0000	***	-7.0397	0.0000	***
wc_acc	-0.0090	0.5372		-0.0120	0.8562		-0.0113	0.8689	
rsst_acc	0.0080	0.3484		0.0230	0.2755		0.0227	0.2491	
ch_rec	1.3071	0.0000	***	1.1311	0.0051	**	1.3365	0.0027	**
ch_inv	0.8998	0.0023	**	0.9244	0.1421		1.0573	0.1034	
soft_assets	1.6666	0.0000	***	1.0998	0.0000	***	1.0763	0.0001	***
ch_cs	-0.0007	0.0828		0.0003	0.8828		0.0009	0.7544	
ch_cm	-0.0005	0.0018	**	-0.0001	0.9075		-0.0001	0.9188	
ch_roa	-0.0009	0.8087		-0.0028	0.9004		-0.0015	0.9373	
ch_fcf	0.0000	0.8409		0.0000	0.7869		0.0000	0.7896	
tax	-0.0119	0.9691		-0.2104	0.4250		-0.2362	0.3706	
ch_emp	0.0007	0.7755		0.0018	0.8195		0.0015	0.8440	
EMP DIFF	-0.0037	0.4002		-0.0044	0.5594		-0.0009	0.9425	
leasedum	0.6594	0.0000	***	0.6618	0.0095	**	0.6544	0.0107	*
oplease	-0.1790	0.1681		-0.3067	0.0299	*	-0.3051	0.0323	*
issue	1.3276	0.0000	***	1.6426	0.0000	***	1.6325	0.0000	***
cff	-0.0416	0.4435		-0.2757	0.0480	*	-0.2507	0.0755	
leverage	-0.0202	0.2566		0.0919	0.0064	**	0.0919	0.0062	**
bm	-0.0006	0.5955		-0.0020	0.8187		-0.0020	0.8136	
ep	0.0006	0.4155		0.0233	0.0715		0.0230	0.0765	
rect_sales							-0.4349	0.4097	
rect_sales_indcomp							0.1310	0.3171	

ch_sales					-0.0201	0.5757
ch_sales_indcomp					-0.0164	0.6623
ch_rect_ch_sales					-0.0273	0.3367
ch_rect_ch_sales_indcomp					-0.0003	0.6133
Misstating firm-years		1,068		328		328
Nonmisstating firm-years		180,628		25,163		25,163
		181,696		25,491		25,491
R-Squared		0.040		0.031		0.032

Panel B: Retail

	Dechow Model - All Ind.			Dechow Model - Retail			Dechow Model PLUS - Retail		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.8690	0.0000	***	-7.1460	0.0000	***	-6.7719	0.0057	**
wc_acc	-0.0090	0.5372		-0.1575	0.6707		-0.2372	0.5385	
rsst_acc	0.0080	0.3484		0.1161	0.7509		0.1910	0.6148	
ch_rec	1.3071	0.0000	***	1.4549	0.2139		2.7648	0.0432	*
ch_inv	0.8998	0.0023	**	-0.2423	0.8200		-0.6180	0.5656	
soft_assets	1.6666	0.0000	***	1.7418	0.0016	**	2.1187	0.0003	***
ch_cs	-0.0007	0.0828		0.0012	0.8814		0.0011	0.8943	
ch_cm	-0.0005	0.0018	**	0.0003	0.9540		0.0022	0.8501	
ch_roa	-0.0009	0.8087		0.0140	0.7203		0.0151	0.7164	
ch_fcf	0.0000	0.8409		0.0001	0.8208		0.0001	0.8515	
tax	-0.0119	0.9691		0.0529	0.9449		0.1645	0.8413	
ch_emp	0.0007	0.7755		0.0182	0.0918		0.0185	0.1578	
EMP DIFF	-0.0037	0.4002		-0.0261	0.0784		-0.0241	0.1623	
leasedum	0.6594	0.0000	***	-0.3883	0.4179		-0.4610	0.3435	
oplease	-0.1790	0.1681		0.1986	0.7363		0.1516	0.8100	
issue	1.3276	0.0000	***	1.5233	0.0120	*	1.4944	0.0137	*
cff	-0.0416	0.4435		0.0916	0.5571		0.1060	0.5326	
leverage	-0.0202	0.2566		-0.0093	0.9165		0.0086	0.9203	
bm	-0.0006	0.5955		-0.0175	0.2584		-0.0179	0.2456	
ep	0.0006	0.4155		0.0081	0.3614		0.0089	0.3002	
rect_sales							-6.9577	0.1195	
rect_sales_indretail							0.2965	0.2644	
ch_sales							-0.0246	0.6332	
ch_sales_indretail							0.0205	0.3536	
sss							0.0067	0.8758	

sss_indretail					0.0121	0.7177	
ch_sss					-0.1136	0.0446	*
ch_sss_indretail					-0.0021	0.0671	
ch_net_sales_sqft					-2.7117	0.7722	
ch_net_sales_sqft_indretail					-0.0123	0.8731	
net_sales_stores					0.0000	0.9471	
net_sales_stores_indretail					0.2699	0.8988	
ch_net_sales_stores					3.8992	0.6324	
ch_net_sales_stores_indretail					-0.0011	0.9901	
Misstating firm-years		1,068		76		76	
Nonmisstating firm-years		180,628		12,831		12,831	
		181,696		12,907		12,907	
R-Squared		0.040		0.029		0.040	

Panel C: Services									
	Dechow Model - All Ind.			Dechow Model - Services			Dechow Model PLUS - Services		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.8690	0.0000	***	-8.2619	0.0000	***	-8.1747	0.0000	***
wc_acc	-0.0090	0.5372		-0.0072	0.8419		-0.0083	0.8048	
rsst_acc	0.0080	0.3484		0.0078	0.7671		0.0090	0.7117	
ch_rec	1.3071	0.0000	***	0.8810	0.0923		0.8780	0.0972	
ch_inv	0.8998	0.0023	**	1.8147	0.0048	**	1.8231	0.0047	**
soft_assets	1.6666	0.0000	***	1.7519	0.0000	***	1.7585	0.0000	***
ch_cs	-0.0007	0.0828		-0.0009	0.0190	*	-0.0009	0.0182	*
ch_cm	-0.0005	0.0018	**	-0.0002	0.7593		-0.0002	0.7495	
ch_roa	-0.0009	0.8087		-0.0007	0.9322		-0.0005	0.9464	
ch_fcf	0.0000	0.8409		0.0000	0.8805		0.0000	0.8453	
tax	-0.0119	0.9691		0.4033	0.3245		0.4117	0.3143	
ch_emp	0.0007	0.7755		-0.0001	0.9820		-0.0001	0.9844	
EMP DIFF	-0.0037	0.4002		-0.0031	0.6925		-0.0031	0.6928	
leasedum	0.6594	0.0000	***	1.0239	0.0036	**	1.0092	0.0043	**
oplease	-0.1790	0.1681		-0.0759	0.8339		-0.0737	0.8396	
issue	1.3276	0.0000	***	1.3849	0.0004	***	1.3725	0.0005	***
cff	-0.0416	0.4435		-0.0097	0.9102		-0.0116	0.8959	
leverage	-0.0202	0.2566		0.0272	0.3181		0.0275	0.3133	
bm	-0.0006	0.5955		-0.0009	0.7205		-0.0009	0.7200	
ep	0.0006	0.4155		0.0041	0.4654		0.0041	0.4660	
rect_sales							-0.1693	0.2639	
rect_sales_indservices							0.0662	0.1845	
Misstating firm-years		1,068			149			149	
Nonmisstating firm-years		180,628			22,263			22,263	

	181,696	22,412	22,412
R-Squared	0.040	0.054	0.055

Table 5
Regression Results by Account: Various Sample Periods

Panel A: Revenue

	Initial Period			Secondary Period			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-8.5156	0.0000	***	-8.5415	0.0000	***	-8.5192	0.0000	***
WC_acc	-0.0201	0.4559		-0.0054	0.8886		-0.0099	0.5957	
rsst_acc	0.0182	0.3305		0.0079	0.7504		0.0078	0.4793	
ch_rec	1.8262	0.0000	***	1.8949	0.0009	***	1.7980	0.0000	***
ch_inv	1.2944	0.0007	***	0.9153	0.3685		1.2141	0.0006	***
soft_assets	2.1755	0.0000	***	2.1266	0.0000	***	2.1631	0.0000	***
ch_cs	-0.0002	0.8933		0.0001	0.9633		-0.0001	0.9480	
ch_cm	0.0000	0.9884		0.0000	0.9740		0.0000	0.9914	
ch_roa	-0.0027	0.5833		-0.0017	0.8906		-0.0008	0.8601	
ch_fcf	0.0000	0.8990		0.0000	0.9032		0.0000	0.8978	
tax	-0.0443	0.9169		0.1714	0.7811		0.0246	0.9510	
ch_emp	0.0003	0.8781		0.0018	0.7429		0.0007	0.8001	
EMP DIFF	-0.0011	0.7534		-0.0047	0.7028		-0.0016	0.6653	
leasedum	-0.5022	0.0004	***	-0.3481	0.3095		-0.4848	0.0002	***
oplease	-0.4308	0.0044	**	-0.0502	0.9019		-0.2687	0.0248	*
issue	1.6002	0.0000	***	1.7150	0.0037	**	1.6112	0.0000	***
cff	-0.0017	0.9702		-0.3071	0.1676		-0.0304	0.6056	
leverage	0.0048	0.6602		-0.0617	0.2869		0.0002	0.9882	
bm	-0.0013	0.5922		-0.0002	0.9001		-0.0008	0.6369	
ep	0.0008	0.4779		0.0004	0.8549		0.0007	0.4907	
Misstating firm-years		452			115			567	
Nonmisstating firm-years		144,747			36,382			181,129	
		145,199			36,497			181,696	

R-Squared	0.053	0.045	0.050

Panel B: Accounts Receivable									
	Initial Period			Secondary Period			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-9.211	0.000	***	-9.430	0.000	***	-9.1151	0.0000	***
WC_acc	-0.004	0.954		-0.003	0.969		-0.0030	0.9416	
rsst_acc	0.006	0.925		0.005	0.913		0.0034	0.9050	
ch_rec	1.775	0.000	***	2.169	0.010	*	1.8135	0.0000	***
ch_inv	1.644	0.003	**	0.485	0.782		1.4556	0.0055	**
soft_assets	2.838	0.000	***	1.762	0.006	**	2.5505	0.0000	***
ch_cs	0.000	0.873		0.000	0.951		-0.0002	0.8863	
ch_cm	0.000	0.863		0.000	0.992		-0.0001	0.9450	
ch_roa	-0.001	0.909		-0.002	0.926		-0.0015	0.8747	
ch_fcf	0.000	0.958		0.000	0.961		0.0000	0.9964	
tax	-0.036	0.959		0.326	0.616		0.1301	0.8330	
ch_emp	0.001	0.899		0.002	0.810		0.0011	0.8377	
EMP DIFF	-0.007	0.418		-0.003	0.855		-0.0061	0.4489	
leasedum	-1.143	0.000	***	-0.478	0.376		-1.0232	0.0001	***
oplease	-0.665	0.001	**	-0.157	0.716		-0.3757	0.0023	**
issue	0.882	0.004	**	1.927	0.058		0.9997	0.0005	***
cff	-0.052	0.659		-0.302	0.377		-0.0520	0.6586	
leverage	0.009	0.466		0.004	0.948		0.0089	0.4962	
bm	-0.001	0.831		0.000	0.932		-0.0005	0.8241	
ep	0.001	0.747		0.000	0.903		0.0006	0.7224	
Misstating firm-years		170			49			219	
Nonmisstating firm-years		145,029			36,448			181,477	
		145,199			36,497			181,696	
R-Squared		0.061			0.036			0.053	

Panel C: Inventory									
	Initial Period			Secondary Period			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-9.5466	0.0000	***	-9.0654	0.0000	***	-9.8284	0.0000	***
WC_acc	-0.0196	0.8345		0.0016	0.9869		-0.0076	0.8793	
rsst_acc	0.0220	0.7359		-0.0029	0.9672		0.0078	0.8321	
ch_rec	0.6142	0.3521		0.7043	0.6398		0.6469	0.2860	
ch_inv	2.7418	0.0000	***	2.3589	0.0257	*	2.6931	0.0000	***
soft_assets	2.8786	0.0000	***	3.3098	0.0003	***	2.9801	0.0000	***
ch_cs	-0.0001	0.9553		-0.0003	0.9224		-0.0001	0.9360	
ch_cm	-0.0005	0.5516		0.0001	0.9358		-0.0003	0.6443	
ch_roa	-0.0029	0.7542		-0.0047	0.7457		-0.0030	0.7537	
ch_fcf	0.0000	0.9729		0.0000	0.9985		0.0000	0.9199	
tax	0.0531	0.9588		0.3066	0.7929		0.1470	0.8664	
ch_emp	0.0003	0.9341		0.0031	0.7764		0.0009	0.8898	
EMP DIFF	-0.0070	0.4914		-0.0097	0.7137		-0.0071	0.4620	
leasedum	-0.8413	0.0064	**	-0.7533	0.3095		-0.7885	0.0055	**
oplease	-0.1339	0.8232		-0.1250	0.8445		-0.1299	0.7614	
issue	0.8706	0.0129	*	15.9879	0.9684		1.0992	0.0016	**
cff	-0.2709	0.3918		-0.9547	0.0771		-0.4921	0.0941	
leverage	0.0073	0.6745		0.0065	0.9237		0.0071	0.6810	
bm	-0.0007	0.8578		-0.0001	0.9522		-0.0006	0.8382	
Ep	0.0007	0.7550		0.0005	0.8962		0.0007	0.7403	
Misstating firm-years		124			31			155	
Nonmisstating firm-years		145,075			36,466			181,541	
		145,199			36,497			181,696	
R-Squared		0.057			0.057			0.059	

Table 6
Regression Results by Industry: Various Sample Periods

Panel A: All Industries									
	Initial Period			Secondary Period			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.8641	0.0000	***	-7.9391	0.0000	***	-7.8690	0.0000	***
wc_acc	-0.0182	0.3755		-0.0046	0.8924		-0.0090	0.5372	
rsst_acc	0.0176	0.1873		0.0060	0.7884		0.0080	0.3484	
ch_rec	1.2916	0.0000	***	1.6562	0.0011	**	1.3071	0.0000	***
ch_inv	1.0549	0.0009	**	0.2083	0.8149		0.8998	0.0023	**
soft_assets	1.7432	0.0000	***	1.4322	0.0000	***	1.6666	0.0000	***
ch_cs	-0.0008	0.0368		-0.0001	0.9096		-0.0007	0.0828	
ch_cm	-0.0007	0.0004	***	-0.0002	0.7845		-0.0005	0.0018	**
ch_roa	-0.0027	0.4335		-0.0019	0.8622		-0.0009	0.8087	
ch_fcf	0.0000	0.8457		0.0000	0.8750		0.0000	0.8409	
tax	-0.1274	0.6460		0.2448	0.5205		-0.0119	0.9691	
ch_emp	0.0004	0.8354		0.0014	0.7413		0.0007	0.7755	
EMP DIFF	-0.0034	0.4805		-0.0049	0.6179		-0.0037	0.4002	
leasedum	0.6373	0.0000	***	0.6801	0.0126	*	0.6594	0.0000	***
oplease	-0.2640	0.1217		-0.0838	0.7718		-0.1790	0.1681	
issue	1.2533	0.0000	***	1.6989	0.0000	***	1.3276	0.0000	***
cff	0.0094	0.7486		-0.4987	0.0109	*	-0.0416	0.4435	
leverage	-0.0020	0.8879		-0.1061	0.0230	*	-0.0202	0.2566	
bm	-0.0008	0.6178		-0.0002	0.8761		-0.0006	0.5955	
ep	0.0007	0.4115		0.0003	0.8275		0.0006	0.4155	
Misstating firm-years		834			234			1,068	
Nonmisstating firm-years		144,365			36,263			180,628	
		145,199			36,497			181,696	
R-Squared		0.042			0.041			0.040	

Panel B: Computer Hardware and Software

	Initial Period			Secondary Period			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.4785	0.0000	***	-5.4039	0.0000	***	-7.0397	0.0000	***
wc_acc	-0.0735	0.5034		0.0173	0.9010		-0.0113	0.8689	
rsst_acc	0.0521	0.1787		0.0597	0.3335		0.0227	0.2491	
ch_rec	1.5880	0.0012	**	-0.0201	0.9871		1.3365	0.0027	**
ch_inv	0.8963	0.1891		2.4508	0.2204		1.0573	0.1034	
soft_assets	1.3450	0.0000	***	0.5822	0.2545		1.0763	0.0001	***
ch_cs	0.0008	0.8094		0.0012	0.9060		0.0009	0.7544	
ch_cm	0.0004	0.9165		-0.0007	0.9307		-0.0001	0.9188	
ch_roa	-0.0105	0.6942		-0.0129	0.7762		-0.0015	0.9373	
ch_fcf	0.0000	0.9867		0.0000	0.7507		0.0000	0.7896	
tax	-0.2667	0.2477		2.8600	0.0586		-0.2362	0.3706	
ch_emp	0.0015	0.8685		-0.0048	0.8461		0.0015	0.8440	
EMP DIFF	-0.0014	0.9170		0.0186	0.8871		-0.0009	0.9425	
leasedum	0.7456	0.0107	*	0.0612	0.9085		0.6544	0.0107	*
oplease	-0.6667	0.0020	**	0.8874	0.4370		-0.3051	0.0323	*
issue	1.6470	0.0001	***	1.7190	0.0233	*	1.6325	0.0000	***
cff	-0.1342	0.3219		-0.9319	0.0155	*	-0.2507	0.0755	
leverage	0.1136	0.0008	***	-0.0388	0.7442		0.0919	0.0062	**
bm	-0.0002	0.9407		-0.0001	0.9623		-0.0020	0.8136	
ep	0.0202	0.1447		0.0927	0.1341		0.0230	0.0765	
rect_sales	0.0423	0.9235		-3.6664	0.0530		-0.4349	0.4097	
rect_sales_indcomp	-0.0048	0.9710		0.8782	0.0300	*	0.1310	0.3171	
ch_sales	-0.0305	0.3832		0.0321	0.8434		-0.0201	0.5757	
ch_sales_indcomp	-0.0025	0.9322		-0.1296	0.4528		-0.0164	0.6623	
ch_rect_ch_sales	-0.0341	0.2674		-0.0464	0.7321		-0.0273	0.3367	
ch_rect_ch_sales_indcomp	0.0006	0.6401		-0.0007	0.3469		-0.0003	0.6133	
Misstating firm-years		244			84			328	
Nonmisstating firm-years		18,969			6,194			25,163	
		19,213			6,278			25,491	

R-Squared	0.039	0.041	0.032

Panel C: Retail									
	Initial Period			Secondary Period			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-6.5899	0.3173		-6.8318	0.0007	***	-6.7719	0.0057	**
wc_acc	-0.6117	0.4682		-0.1767	0.8118		-0.2372	0.5385	
rsst_acc	0.4230	0.3428		0.1417	0.8449		0.1910	0.6148	
ch_rec	3.4247	0.0339	*	2.7788	0.5915		2.7648	0.0432	*
ch_inv	0.2973	0.8172		-2.4167	0.3142		-0.6180	0.5656	
soft_assets	1.6847	0.0137	*	3.9120	0.0035	**	2.1187	0.0003	***
ch_cs	0.0006	0.9283		0.0739	0.8158		0.0011	0.8943	
ch_cm	0.0010	0.9384		0.0401	0.7142		0.0022	0.8501	
ch_roa	-0.7147	0.0617		0.0129	0.8398		0.0151	0.7164	
ch_fcf	0.0228	0.1405		0.0000	0.9023		0.0001	0.8515	
tax	-0.2928	0.7232		3.7818	0.5915		0.1645	0.8413	
ch_emp	0.0162	0.2258		-0.0048	0.8394		0.0185	0.1578	
EMP DIFF	-0.0198	0.2654		0.2128	0.4513		-0.0241	0.1623	
leasedum	-0.5450	0.3211		-0.0191	0.9872		-0.4610	0.3435	
oplease	0.3113	0.6720		-0.1633	0.8593		0.1516	0.8100	
issue	1.2595	0.0445	*	16.3932	0.9944		1.4944	0.0137	*
cff	0.3631	0.3242		0.0551	0.8761		0.1060	0.5326	
leverage	-0.0029	0.9800		0.0037	0.9687		0.0086	0.9203	
bm	-0.0118	0.5465		-0.0572	0.5443		-0.0179	0.2456	
ep	0.0104	0.2727		0.2224	0.4239		0.0089	0.3002	
rect_sales	-10.3373	0.0577		-12.5698	0.5571		-6.9577	0.1195	
rect_sales_indretail	0.5808	0.0762		0.0663	0.9490		0.2965	0.2644	
ch_sales	-0.0324	0.6916		-0.0003	0.9993		-0.0246	0.6332	
ch_sales_indretail	0.0252	0.4404		-0.1013	0.5702		0.0205	0.3536	
sss	-0.0886	0.3670		0.0518	0.4134		0.0067	0.8758	
sss_indretail	0.0951	0.3623		0.0107	0.8048		0.0121	0.7177	
ch_sss	-0.2719	0.0007	***	0.0465	0.6141		-0.1136	0.0446	*
ch_sss_indretail	-0.0651	0.2260		-0.0010	0.4032		-0.0021	0.0671	
ch_net_sales_sqft	1.6252	0.9382		-2.1285	0.8741		-2.7117	0.7722	
ch_net_sales_sqft_indretail	-0.4148	0.2885		-0.0030	0.9680		-0.0123	0.8731	

net_sales_stores	0.0000	0.9325	0.0000	0.8739	0.0000	0.9471
net_sales_stores_indretail	-0.3200	0.9580	0.0845	0.8757	0.2699	0.8988
ch_net_sales_stores	8.6217	0.4096	-0.8218	0.9564	3.8992	0.6324
ch_net_sales_stores_indretail	0.1595	0.4421	-0.0020	0.9870	-0.0011	0.9901
Misstating firm-years		59		17		76
Nonmisstating firm-years		10,666		2,165		12,831
		10,725		2,182		12,907
R-Squared		0.057		0.106		0.040

Panel D: Services									
	Initial Period			Secondary Period			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.9681	0.0000	***	-7.3652	0.0000	***	-8.1747	0.0000	***
wc_acc	-0.0114	0.8720		0.0206	0.7700		-0.0083	0.8048	
rsst_acc	0.0158	0.6437		-0.0167	0.7065		0.0090	0.7117	
ch_rec	0.8196	0.1600		1.8674	0.1641		0.8780	0.0972	
ch_inv	2.0093	0.0082	**	1.9682	0.1626		1.8231	0.0047	**
soft_assets	1.7111	0.0000	***	1.7161	0.0125	*	1.7585	0.0000	***
ch_cs	-0.0009	0.0181	*	-0.0053	0.6159		-0.0009	0.0182	*
ch_cm	-0.0003	0.6863		-0.0001	0.9922		-0.0002	0.7495	
ch_roa	-0.0056	0.8095		-0.0030	0.8534		-0.0005	0.9464	
ch_fcf	0.0000	0.8548		0.0000	0.9950		0.0000	0.8453	
tax	0.5836	0.1664		-0.5733	0.8405		0.4117	0.3143	
ch_emp	-0.0004	0.8864		0.0005	0.9532		-0.0001	0.9844	
EMP DIFF	-0.0032	0.7264		-0.0035	0.7959		-0.0031	0.6928	
leasedum	0.7964	0.0277	*	xxxx	xxxx		1.0092	0.0043	**
oplease	0.0434	0.9062		-0.4558	0.5942		-0.0737	0.8396	
issue	1.2706	0.0029	**	2.1036	0.0498	*	1.3725	0.0005	***
cff	0.0182	0.7612		-0.7090	0.1399		-0.0116	0.8959	
leverage	0.0308	0.2583		-0.0086	0.9125		0.0275	0.3133	
bm	-0.0008	0.7765		-0.0013	0.8098		-0.0009	0.7200	
ep	0.0037	0.5617		0.0111	0.6528		0.0041	0.4660	
rect_sales	-0.0864	0.6936		-1.1444	0.3400		-0.1693	0.2639	
rect_sales_indservices	0.0366	0.6569		0.2865	0.2698		0.0662	0.1845	
Misstating firm-years		109			40			149	
Nonmisstating firm-years		17,872			4,391			22,263	
		17,981			4,431			22,412	
R-Squared		0.051			0.066			0.055	

Note: For the variable leasedum, “xxxx” indicates that estimation failed when the model included this variable. To allow the regression to run, I removed the leasedum variable.

Table 7
Regression Results by Industry: Decade-Based Time Periods

Panel A: All Industries												
	1980 - 1989			1990 - 1999			2000 - 2008			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.6769	0.0000	***	-7.9239	0.0000	***	-7.7409	0.0000	***	-7.9005	0.0000	***
ch_rec	2.0088	0.0001	***	1.5411	0.0000	***	1.3178	0.0002	***	1.2635	0.0000	***
ch_inv	1.6706	0.0031	**	1.2204	0.0095	**	0.9624	0.0664		0.8679	0.0030	**
soft_assets	1.9193	0.0000	***	1.6426	0.0000	***	1.4858	0.0000	***	1.6889	0.0000	***
ch_cs							-0.0009	0.0651		-0.0007	0.0812	
ch_cm				-0.0007	0.0003	***				-0.0005	0.0018	**
tax	1.8343	0.0381	*									
leasedum				0.7936	0.0000	***	0.5805	0.0005	***	0.6638	0.0000	***
oplease							-0.2243	0.0776				
issue	0.6189	0.0208	*	1.1826	0.0000	***	1.7981	0.0000	***	1.3139	0.0000	***
cff							-0.4089	0.0010	**			
Misstating firm-years		148			404			516			1,068	
Nonmisstating firm-yrs		54,447			68,345			57,836			180,628	
		54,595			68,749			58,352			181,696	
R-Squared		0.041			0.044			0.040			0.040	

Panel B: Computer Hardware and Software												
	1980 - 1989			1990 - 1999			2000 - 2008			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-6.8799	0.0000	***	-6.2734	0.0000	***	-5.8441	0.0000	***	-5.8806	0.0000	***
ch_rec	1.7596	0.0405	*	1.5612	0.0042	**				1.0924	0.0035	**
ch_inv				2.3002	0.0116	*				1.0762	0.0730	
soft_assets	3.5047	0.0016	**	1.0318	0.0120	*	0.9415	0.0056	**	1.0723	0.0000	***
leasedum	-0.8707	0.0319	*	1.2373	0.0153	*	1.3482	0.0096	**	0.8913	0.0004	***
oplease							-0.2880	0.0779		-0.3162	0.0251	*
cff							-0.4422	0.0400	*			
leverage				0.0885	0.0187	*				0.0869	0.0059	**
Misstating firm-years		29			128			171			328	
Nonmisstating firm-yrs		5,192			9,763			10,208			25,163	
		5,221			9,891			10,379			25,491	
R-Squared		0.058			0.032			0.019			0.018	

Panel C: Retail

	1980 – 1989			1990 - 1999			2000 - 2008		All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	Coefficient	p-value	
Intercept	-6.7538	0.0000	***	-6.7108	0.0000	***	-21.203	0.991	-7.5584	0.0000	***
ch_rec	7.0520	0.0026	**								
soft_assets				2.6953	0.0025	**			1.7980	0.0008	***
tax	-65.6836	0.0000	***								
ch_emp				0.0318	0.0254	*			0.0205	0.0509	
EMP DIFF				-0.0901	0.0038	**			-0.0288	0.0500	
oplease	2.0498	0.0729									
issue							16.744	0.993	1.4697	0.0128	*
bm	-0.2037	0.0035	**								
ch_sales_indretail				0.0209	0.0611						
ch_sss									-0.1077	0.0519	
ch_sss_indretail									-0.0020	0.0765	
Misstating firm-years		9			31			36		76	
Nonmisstating firm-yrs		4,177			5,114			3,540		12,831	
		4,186			5,145			3,576		12,907	
R-Squared		0.263			0.043			0.024		0.028	

Panel D: Services												
	1980 – 1989			1990 - 1999			2000 - 2008			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-6.2333	0.0000	***	-7.8473	0.0000	***	-9.3545	0.0000	***	-8.2240	0.0000	***
ch_rec	2.3867	0.0424	*	1.2895	0.0620							
ch_inv	2.5703	0.0651					2.1147	0.0030	**	1.8143	0.0029	**
soft_assets				1.7571	0.0030	**	1.3945	0.0021	**	1.7671	0.0000	***
ch_cs										-0.0009	0.0139	*
tax	8.8924	0.0002	***									
leasedum							2.3936	0.0177	*	1.0070	0.0038	**
issue				1.7279	0.0169	*	1.9282	0.0081	**	1.4266	0.0003	***
Misstating firm-years		17			49			83			149	
Nonmisstating firm-years		6,209			8,794			7,260			22,263	
		6,226			8,843			7,343			22,412	
R-Squared		0.063			0.045			0.067			0.052	

Table 8
Regression Results by Account: Decade-Based Time Periods

Panel A: All Accounts

	1980 - 1989			1990 - 1999			2000 - 2008			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-7.6769	0.0000	***	-7.9239	0.0000	***	-7.7409	0.0000	***	-7.9005	0.0000	***
ch_rec	2.0088	0.0001	***	1.5411	0.0000	***	1.3178	0.0002	***	1.2635	0.0000	***
ch_inv	1.6706	0.0031	**	1.2204	0.0095	**	0.9624	0.0664		0.8679	0.0030	**
soft_assets	1.9193	0.0000	***	1.6426	0.0000	***	1.4858	0.0000	***	1.6889	0.0000	***
ch_cs							-0.0009	0.0651		-0.0007	0.0812	
ch_cm				-0.0007	0.0003	***				-0.0005	0.0018	**
tax	1.8343	0.0381	*									
leasedum				0.7936	0.0000	***	0.5805	0.0005	***	0.6638	0.0000	***
oplease							-0.2243	0.0776				
issue	0.6189	0.0208	*	1.1826	0.0000	***	1.7981	0.0000	***	1.3139	0.0000	***
cff							-0.4089	0.0010	**			
Misstating firm-years		148			404			516			1,068	
Nonmisstating firm-yrs		54,447			68,345			57,836			180,628	
		54,595			68,749			58,352			181,696	
R-Squared		0.041			0.044			0.040			0.040	

Panel B: Revenue

	1980 - 1989			1990 - 1999			2000 - 2008			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-8.3564	0.0000	***	-9.4985	0.0000	***	-8.7613	0.0000	***	-9.0105	0.0000	***
ch_rec	2.5898	0.0000	***	2.1983	0.0000	***	1.5686	0.0001	***	1.7604	0.0000	***
ch_inv	1.7560	0.0095	**	1.7983	0.0012	**				1.2023	0.0006	***
soft_assets	1.9272	0.0001	***	1.9011	0.0000	***	2.2681	0.0000	***	2.1652	0.0000	***
tax	2.2024	0.0172	*									
leasedum	-0.3888	0.0872		0.9575	0.0004	***	0.4017	0.0672		0.4905	0.0002	***
oplease							-0.2608	0.0463	*	-0.2672	0.0205	*
issue	1.0533	0.0088	**	1.7311	0.0000	***	1.7966	0.0000	***	1.6083	0.0000	***
cff							-0.2852	0.0606				
Misstating firm-years		91			211			265			567	
Nonmisstating firm-yrs		54,504			68,538			58,087			181,129	
		54,595			68,749			58,352			181,696	
R-Squared		0.053			0.061			0.047			0.050	

Panel C: Accounts Receivable

	1980 - 1989			1990 - 1999			2000 - 2008			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-8.6062	0.0000	***	-10.9017	0.0000	***	-9.4663	0.0000	***	-10.1416	0.0000	***
ch_rec	2.7340	0.0082	**	2.1587	0.0000	***	1.3035	0.0238	*	1.7574	0.0000	***
ch_inv	3.1654	0.0008	***	2.2027	0.0029	**				1.4249	0.0060	**
soft_assets				3.3563	0.0000	***	2.1824	0.0000	***	2.5597	0.0000	***
leasedum	0.9017	0.0989		1.3493	0.0093	**	0.6439	0.0692		1.0259	0.0001	***
oplease							-0.3128	0.0083	**	-0.3619	0.0020	**
issue				0.7220	0.0917		1.4280	0.0053	**	1.0000	0.0005	***
Misstating firm-years		26			81			112			219	
Nonmisstating firm-yrs		54,569			68,668			58,240			181,477	
		54,595			68,749			58,352			181,696	
R-Squared		0.046			0.081			0.037			0.053	

Panel D: Inventory												
	1980 - 1989			1990 - 1999			2000 - 2008			All Years		
Variable	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value	
Intercept	-10.6788	0.0000	***	-10.4771	0.0000	***	-9.9863	0.0000	***	-10.6626	0.0000	***
ch_inv	2.6892	0.0010	**	2.4521	0.0134	*	3.1853	0.0000	***	2.5406	0.0000	***
soft_assets	3.3592	0.0004	**	2.7439	0.0001	***	3.0271	0.0000	***	3.0515	0.0000	***
leasedum				1.6556	0.0222	*				0.8144	0.0040	**
issue	1.2119	0.0985					1.5824	0.0090	**	1.0647	0.0021	**
cff							-0.7821	0.0080	**			
Misstating firm-years		31			49			75			155	
Nonmisstating firm-yrs		54,564			68,700			58,277			181,541	
		54,595			68,749			58,352			181,696	
R-Squared		0.061			0.048			0.064			0.057	

Appendix A
Variable Calculations

Variable	Abbreviation	Pred sign	Calculation
Misstatement flag	misstate	N/A	1 if misstatement firm-year
Accruals quality related variables			
WC accruals	wc acc	+	((Δ current assets - Δ cash) - (Δ current liabilities - Δ debt in current liabilities - Δ taxes payable)) / average total assets
RSST accruals	rsst acc	+	(Δ WC accruals + Δ ((total assets - current assets - investments and advances) - (total liabilities - current liabilities - long term debt)) + Δ ((short term investments + long term investments) - (long term debt + debt in current liabilitites + preferred stock))) / average total assets
Change in receivables	ch rec	+	Δ accounts receivable / average total assets
Change in inventory	ch inv	+	Δ inventory / average total assets
% Soft assets	soft assets	-	(total assets - PP&E - cash) / total assets
Performance variables			
% Change in cash sales	ch cs	-	((sales - Δ accounts receivable)$_t$ - (sales - Δ accounts receivable)$_{t-n}$) / (sales - Δ accounts receivable)$_{t-n}$
% Change in cash margin	ch cm	-	(1 - ((cost of goods sold - Δ inventory + Δ accounts payable) / ((sales - Δ accounts receivable))$_t$ - (1 - ((cost of goods sold - Δ inventory + Δ accounts payable) / ((sales - Δ accounts receivable)))$_{t-n}$) / (1 - ((cost of goods sold - Δ inventory + Δ accounts payable) / ((sales - Δ accounts receivable))$_{t-n}$
Change in return on assets	ch roa	+	(earnings / average total assets)$_t$ - (earnings / average total assets)$_{t-n}$
Change in free cash flows	ch fcf	-	Δ (earnings - RSST accruals) / average total assets
Deferred tax expense	tax	+	deferred tax expense$_t$ / total assets$_{t-n}$

Nonfinancial variables			
Abnormal change in employees	ch_emp	-	$(((\# \text{ employees})_t - (\# \text{ employees})_{t-n}) / (\# \text{ employees})_{t-n}) - (((\text{total assets})_t - (\text{total assets})_{t-n}) / (\text{total assets})_{t-n})$
Abnormal change in employees	EMP DIFF	-	$(((\text{revenue})_t - (\text{revenues})_{t-n}) / (\text{revenue})_{t-n}) - (((\# \text{ employees})_t - (\# \text{ employees})_{t-n}) / (\# \text{ employees})_{t-n})$
Off-balance-sheet variables			
Existence of operating leases	leasedum	+	1 if future operating lease obligations > 0
Change in operating leases	oplease	+	(Δ Σ(present value of future operating lease obligations)) / average total assets
Market-related variables			
Issuance of securities	issue	+	1 if securities issued during the year
New financing raised	cff	+	amount of financing raised / average total assets
Leverage	leverage	+	long-term debt / total assets
Book to market	bm	-	equity / market value
Earnings to price	ep	-	earnings / market value
New variables – all industries			
Receivables to total sales	rect sales	+	total receivables / sales
Receivables to total sales, compared to industry	rect sales ind	+	(rect sales (firm-year) - rect sales (industry-year)) / rect sales (industry-year)
Computer hardware and software variables			
% Change in total sales	ch sales	+	$(\text{sales}_t - \text{sales}_{t-1}) / \text{sales}_{t-1}$
% Change in total sales, compared to industry	ch sales ind	+	(ch sales (firm-year) - ch sales (industry-year)) / ch sales (industry-year)
% Change in receivables to % change in sales	ch rect ch sales	+	$((\text{total receivables}_t - \text{total receivables}_{t-1}) / \text{total receivables}_{t-1}) - ((\text{sales}_t - \text{sales}_{t-1}) / \text{sales}_{t-1})$
% Change in receivables to % change in sales, compared to industry	ch rect ch sales ind	+	(ch rec ch sales (firm-year) - ch rec ch sales (industry-year)) / ch rec ch sales (industry-year)

Retail variables			
% Change in total sales	ch sales	+	$(sales_t - sales_{t-1}) / sales_{t-1}$
% Change in total sales, compared to industry	ch sales ind	+	ch sales (firm-year) - ch sales (industry-year)
Same store sales	sss	+	Data obtained via COMPUSTAT
Same store sales, compared to industry	sss ind	+	sss (firm-year) - sss (industry-year)
% Change in same store sales	ch sss	+	$(sss_t - sss_{t-1}) / sss_{t-1}$
% Change in same store sales, compared to industry	ch sss ind	+	ch sss (firm-year) - ch sss (industry-year)
% Change in sales to square feet retail space	ch net sales sqft	+	$((\text{net sales} / \text{sqft})_t - (\text{net sales} / \text{sqft})_{t-1}) / (\text{net sales} / \text{sqft})_{t-1}$ [net sales / sqft provided by COMPUSTAT]
% Change in sales to square feet retail space, compared to industry	ch net sales sqft ind	+	ch net sales sqft (firm-year) - ch net sales sqft (industry-year)
% Change in sales to number of retail stores	ch net sales stores	+	$((\text{net sales} / \text{average retail stores})_t - (\text{net sales} / \text{average retail stores})_{t-1}) / (\text{net sales} / \text{average retail stores})_{t-1}$
% Change in sales to number of retail stores, compared to industry	ch sales stores ind	+	ch net sales stores (firm-year) - ch net sales stores (industry-year)

REFERENCES

Abbott, L.J., S. Parker, and G.F. Peters. 2004. Audit Committee Characteristics and Restatements. *Auditing: A Journal of Practice & Theory* 23: 69-88.

Aghghaleh, S.F., Z.M. Mohamed, and M.M. Rahmat. 2016. Detecting Financial Statement Frauds in Malaysia: Comparing the Abilities of Beneish and Dechow Models. *Asian Journal of Accounting and Governance* 7: 57-65.

Altamuro, J., A. L. Beatty, and J. Weber. 2005. The Effects of Accelerated Revenue Recognition on Earnings Management and Earnings Informativeness: Evidence from SEC Staff Accounting Bulletin No. 101. *The Accounting Review* 80: 373-401.

Altman, E.I. 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance* 23 (4): 589-609.

American Institute of Certified Public Accountants (AICPA). 1997. *Considerations of Fraud in a Financial Statement Audit* Statement on Auditing Standards No. 82. New York, NY: AICPA.

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

Bayley, L., & S.L. Taylor. 2007. Identifying Earnings Overstatements: A Practical Test. *Working paper.*

Beasley, M.S. 1996. An Empirical Analysis of the Relation Between the Board of Director Composition and Financial Statement Fraud. *The Accounting Review* 71: 443-465.

Beasley, M.S., J.V. Carcello, and D.R. Hermanson. 1999. Fraudulent Financial Reporting: 1987-1997. An Analysis of US Public Companies. *Committee of Sponsoring Organizations of the Treadway Commission (COSO)*

Beasley, M.S., J.V. Carcello, D.R. Hermanson, and T.L. Neal. 2010. Fraudulent Financial Reporting: 1998-2007. An Analysis of US Public Companies. *Committee of Sponsoring Organizations of the Treadway Commission (COSO)*

Bell, T.B., and J.V. Carcello. 2000. A Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting. *Auditing: A Journal of Practice & Theory* 19 (1): 169-184.

Beneish, M.D. 1999. The Detection of Earnings Manipulation. *Financial Analysts Journal* 55 (5): 24-36.

Bonner, S.E., Z-V. Palmrose, and S.M. Young. 1998. Fraud Type and Auditor Litigation: An Analysis of SEC Accounting and Auditing Enforcement Releases. *The Accounting Review* 73 (4): 503-532.

Brazel, J.F., K.L. Jones, and M.F. Zimbelman. 2009. Using Nonfinancial Measures to Assess Fraud Risk. *Journal of Accounting Research* 47 (5): 1135-1166.

Brazel, J.F., K.L. Jones, J. Thayer, and R.C. Warne. 2015. Understanding Investor Perceptions of Financial Statement Fraud and their Use of Red Flags: Evidence from the Field. *Review of Accounting Studies* 20: 1373-1406.

Callen, J.L., S.W. Robb, and D. Segal. 2008. Revenue Manipulation and Restatements by Loss Firms. *Auditing: A Journal of Practice & Theory* 27: 1-29.

Cohen, J.R., L.L. Holder-Webb, L. Nath, and D. Wood. 2012. Corporate Reporting of Nonfinancial Leading Indicators of Economic Performance and Sustainability. *Accounting Horizons* 26 (1): 65-90.

Dechow, P.M., W. Ge, C.R. Larson, and R.G. Sloan. 2011. Predicting Material Accounting Misstatements. *Contemporary Accounting Research* 28: 17-82.

Dechow, P.M., S. Richardson, and I. Tuna. 2003. Why are Earnings Kinky? An Examination of the Earnings Management Explanation. *Review of Accounting Studies* 8: 355-384.

Dechow, P.M. and A. Sweeney. 1995. Detecting Earnings Management. *The Accounting Review* 70: 193-225.

Farber, D. 2005. Restoring Trust After Fraud: Does Corporate Governance Matter? *The Accounting Review* 80: 539-561.

Feroz, E.H., K. Park, and V.S. Pastena. 1991. The Financial and Market Effects of the SEC's Accounting and Auditing Enforcement Releases. *Journal of Accounting Research* 29: 107-142.

Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder Litigation and Corporate Disclosures. *Journal of Accounting Research* 32 (Autumn): 137-164.

Hogan, C.E., Z. Rezaee, R.A. Riley, Jr., and U.K. Velury. 2008. Financial Statement Fraud: Insights from the Academic Literature. *Auditing: A Journal of Practice & Theory* 27 (2): 231-252.

Jones, J. 1991. Earnings Management during Import Relief Investigations. *Journal of Accounting Research* 29: 193-228.

Kaminski, K.A., and T.S. Wetzel. 2004. Financial Ratios and Fraud: An Exploratory Study using Chaos Theory. *Journal of Forensic Accounting* V: 147-172.

Kaminski, K.A., T.S. Wetzel, and L. Guan. 2004. Can Financial Ratios Detect Fraudulent Financial Reporting? *Managerial Auditing Journal* 19 (1): 15-28.

Marquardt, C.A., and C.I. Wiedman. 2004. How Are Earnings Managed? An Examination of Specific Accruals. *Contemporary Accounting Research* 21: 461-491.

Palmrose, Z-V. 1988. An Analysis of Auditor Litigation and Audit Service Quality. *The Accounting Review* 63 (January): 55-73.

Public Company Accounting Oversight Board (PCAOB). 2010. *Identifying and Assessing Risks of Material Misstatement* Accounting Standard No. 2110. Washington, DC: PCAOB

Rezaee, Z. 2005. Causes, Consequences, and Deterrence of Financial Statement Fraud. *Critical Perspectives on Accounting* 16: 277-298.

Securities and Exchange Commission (SEC). 1999. *Revenue Recognition.* Staff Accounting Bulletin No. 101. Washington D.C.: Government Printing Office.

Trompeter, G., T. Carpenter, N. Desai, K. Jones, and R. Riley. 2013. A Synthesis of Fraud Related Research. *Auditing: A Journal of Practice and Theory* 32: 287-321.

Zmijewski, M.E. 1984. Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research* 22: 59-82.

VITA

Emily D. Cokeley

EDUCATION

PhD (Business Administration – Accounting), Syracuse University, December 2018
BS (Accounting), State University of New York College at Brockport, May 2006
BS (Psychology), Bethany College, May 2003

PROFESSIONAL EXPERIENCE

Accountant, Excellus BCBS; Rochester, NY 2010 – 2011
Audit Senior, EFP Rotenberg; Rochester, NY 2007 – 2010
Audit Staff, KPMG; Rochester, NY 2006 – 2007

TEACHING EXPERIENCE

Lecturer, Rochester Institute of Technology; Rochester, NY 2014 – Present
Instructor, Syracuse University; Syracuse, NY 2014

AWARDS AND HONORS

AICPA Accounting Doctoral Scholars Program (ADS)

PROFESSIONAL MEMBERSHIPS

Active New York State CPA License
American Accounting Association
American Institute of CPAs
New York State Society of CPAs