THE RELATIONSHIP BETWEEN AN INDEPENDENT AUDIT AND FINANCIAL REPORTING QUALITY: EVIDENCE FROM SMALL PRIVATE COMMERCIAL BANKS

By David Y. Chan

A dissertation submitted to the

Graduate School-Newark

Rutgers, The State University of New Jersey

in partial fulfillment of requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Dr. Alexander Kogan

and approved by

____
Dr. Alexander Kogan

____
Dr. Miklos A. Vasarhelyi

____
Dr. Michael G. Alles

____
Dr. Marilyn Prosch

____
Dr. Divya Anantharaman

Newark, New Jersey
May 2012

ABSTRACT

The Relationship between an Independent Audit and Financial Reporting Quality:

Evidence from Small Private Commercial Banks

By David Y. Chan

Dissertation Director: Dr. Alexander Kogan

The study starts by examining whether the decision to have an independent audit is systematic and endogenous. Private commercial banks with under $500 million in total assets are regulatory exempt from an independent audit requirement. However, approximately 56% of these small private commercial banks voluntarily decided to have an independent audit. Utilizing a set of machine learning algorithms, I find that the decision to have an independent audit is systematically determinable and is endogenously determined given a set of bank characteristics. I also find that bank size, profitability, growth, complexity of operations, and type of ownership may influence the decision to have an independent audit at small private commercial banks.

In the main part of the study, I analyze whether an independent audit improves financial reporting quality. Financial reporting quality is quantitatively measured by material accuracy, conservative recognition of probable losses, and the magnitude of discretionary accruals. Based on these quantitative measures, I find that having an independent audit does not improve financial reporting quality. More specifically, the

results in this study indicate that independently audited small private commercial banks had a higher likelihood of having a restatement, were less conservative in recognizing probable loan losses, and had higher magnitudes of discretionary accruals. Collectively, the results in this study provide rigorous and substantial evidence that the quality of financial reporting may not increase by having an independent audit and may not support the benefits of procuring an independent audit absent regulatory requirements. Moreover, the study implicates the quality of audits being performed at small private commercial banks and perhaps suggests that the audit methodology used by external auditors has to be reconsidered.

ACKNOWLEDGEMENTS

I would like to express here my deepest gratitude to my dissertation committee members. First, I would like to specially thank Dr. Alexander Kogan, my dissertation chair, for all his encouragement, guidance, and support during my doctoral studies and the dissertation process. I am forever indebted to Dr. Kogan's mentorship. Second, I would like to thank Dr. Miklos Vasarhelyi for giving me the opportunity to pursue a Ph.D. in Accounting and Information Systems at Rutgers University. Third, I would like to thank Dr. Michael Alles for teaching me to look at scholarly work using a critical mindset. And fourth, I would like to thank Dr. Marilyn Prosch and Dr. Divya Anantharaman for all their constructive and helpful comments on my dissertation.

I would also like to take this opportunity to thank my parents, Anna and Raymond Chan, and my brother, Alan Chan. I could not have completed my doctoral studies and this dissertation without their love and support.

TABLE OF CONTENTS

TABLES, FIGURES, AND MODELS

TABLES

FIGURES

MODELS

Chapter 1

1. INTRODUCTION

Private companies in general are not mandated to have an independent audit. In contrast, public companies and some private companies in highly regulated industries such as banking and utilities have regulatory audit requirements. The imposition of a mandatory audit requirement for all types of companies can be considered an intrusive regulatory and economic burden (Keasey et al. 1988). However, absent regulatory requirements, there is substantial evidence that an independent audit has a perceived value and benefits that exceed the economic burden. Prior to the enactment of the Securities Exchange Act of 1934[1], approximately 94% of public companies on the New York Stock Exchange (NYSE) were already audited by independent auditors (Wallace 1980). As for private companies, (Collis 2010; Collis et al. 2004; Seow 2001; Rennie et al. 2003) find many small private companies who are exempt from regulatory requirements still opted to have an independent audit.

The demand for an independent audit may result from the need for a monitoring mechanism, quality financial information, and insurance against reliance damages by stakeholders (Wallace 1980). The stakeholders or principals can include owners, investors, creditors, customers, taxpayers, and even regulators. In agency theory (Watts and Zimmerman 1983; Jensen and Meckling 1976; Flint 1988), the principal has an

[1] Mandated the preparation and certification of financial statements by public companies.

inherent interest in an independent audit because it can reduce opportunistic behavior and reporting by the agent. An independent audit can reduce agent bias in reporting and improve the quality of financial reporting information (Kinney Jr and Martin 1994). Furthermore, an independent audit reduces the level of information uncertainty for outsiders by having an independent third party attest to the material recording accuracy of the company's business transactions and accounting estimates. The objective of an independent audit is to express an opinion on whether the financial statements are free from material misstatements (AU Section 110). A misstatement is material if it can change or influence the judgment of a reasonable person (AU Section 312).

Arguably, the premise of the demand for an independent audit revolves around the demand for quality financial information. Prior literature has focused on the effect that auditor characteristics or type of auditing firm has on financial reporting quality at public companies. For example, (DeAngelo 1981) finds that accounting firm size may be a good proxy for audit quality and (Turner and Sennetti 2001) find that companies audited by the Big 6 are less likely to have restatements. Other researchers study the effect of auditor specialization and tenure on audit and financial reporting quality. The big audit firms typically have specialized industry practices. Auditors who have industry specialty expertise can perform better quality audits (Solomon et al. 1999). Additionally, a reoccurring auditor can perform a higher quality audit and thus improve financial reporting quality due to continual experience with the client being audited (Johnson et al. 2002). Collectively, these characteristics and attributes of auditors may enhance the overall level of financial reporting quality. However, to the best of our knowledge, there has been anecdotal and circumstantial evidence but no rigorous empirical studies showing the correlation between an independent audit and financial reporting quality. By definition of an independent audit, it is un-debatable that an audit improves financial reporting quality. In other words, the financial reporting information of audited companies should be of higher quality than unaudited companies. The amount of research in this is area has been limited because of the lack of financial reporting data on private companies with no regulatory audit requirement.

In this study, we specifically examine and gain insight on the relationship between an independent audit and financial reporting quality in the small private commercial bank setting. Private commercial banks (also referred to subsequently as banks) are regulated by the Federal Deposit Insurance Corporation (FDIC). Unlike public companies, private commercial banks are not all require by FDIC to have an independent audit. Private commercial banks with under $500 million in total assets are regulatorily exempt from having an independent audit[2,3]. The non-mandate could be rationalized by the potential regulatory burden that an independent audit can have on these smaller banks. Another potential explanation of only requiring independent audits of banks with above $500 million in total assets is that larger banks may pose a

---

[2] FDI Act 12 C.F.R. Part 363

[3] 3 Exceptions: 1) Thrifts, regardless of size, with a composite safety and soundness CAMELS rating of 3,4,5; 2) Holding companies which control insured financial institution subsidiary(ies) with aggregate consolidated assets of $500 million or more; 3) Any other entity for which the OTS determines an audit is required for safety and soundness reasons; (OTS Audit Rule 12 CFR 562.4) and 4) Small banks part of a public holding company.

greater systemic risk to the financial system. Thus, the incentive is greater for more accurate reporting and monitoring of larger banks. This was exhibited by the recent financial crisis where the government only bailed out larger banks under the notion that they were too big to fail. Absent regulatory requirement, there still appears to be a perceived value of an independent audit by smaller banks. In this study, we find approximately 56% of small private commercial banks decided to have an independent audit absent a regulatory requirement. An independent audit can also serve as a signal to external parties that the company's financial reporting is of higher quality when compared with unaudited companies (Wallace 1980). For banks, having an independent audit can be used to signal to regulators that its financial reporting is of higher quality and perhaps may lower the probability of triggering a potential supervisory examination.

Commercial banks are an integral part of our economy. They provide an essential public financial intermediary service between borrowers and depositors. Evident by the recent financial crisis, the failure of banks can be catastrophically damaging to economic production. Therefore, it is in the interest of the public to have these institutions carefully monitored. The FDIC is mandated by the government to independently maintain stability and public confidence in the banking system. The FDIC uses Call Report filings and supervisory examinations as a means to monitor banks. Call Reports are similar to the financial statement filings by public companies with the Securities and Exchange Commission (SEC). For the FDIC to effectively supervise and examine banks, timely and accurate financial reporting is essential. Material misstatements or misleading information on the Call Reports undermined the ability of Regulators to effectively monitor and supervise banks. According to (Lindo 2007), the most common causes of material misstatements at banking organizations were relating to the misapplication of accounting principles and calculation errors. Although not all banks are required to have an independent audit, an audit should have a significant role in reducing material misstatements and the misapplication of accounting principles.

We start our study by examining whether the decision to have an audit is a systematic and endogenous decision based on a set of company characteristics. We test this theory by training and testing six popular machine learning classification algorithms. To the best of our knowledge, this is the first study to utilize machine learning algorithms to test whether the decision to have an independent audit is systematic and endogenous. Prior accounting and auditing literature has looked at the use of machine learning algorithms to predict the likelihood that a company is a going concern, going bankrupt, and committing fraud. We find that four of the six machine learning algorithms used were able to learn and then classify with reasonable accuracy whether a bank is likely to be audited utilizing only bank characteristic variables (Accuracy Rates: SimpleLogistics (72.23%), SMO (71.81%), JRIP (71.92%), and RandomForest (72.45%)). The reasonable accuracy rates of these machine learning algorithms suggest that the decision to have an audit may be systematically determinable and is endogenously and optimally determined given company characteristics. We also find specifically that bank size, profitability, growth, complexity of operations, and type of ownership may influence the decision to have an independent audit.

In the main part of the study, we analyze whether an independent audit improves financial reporting quality. We measure financial reporting quality quantitatively using material accuracy, conservative recognition of losses, and the conservative use of discretionary accruals in reporting. First, restatements are used as a proxy for material accuracy in financial reporting. A restatement is required when a material misstatement is discovered subsequently on issued financial statements. Intuition and anecdotal evidence would suggest that consistently audited banks should have fewer restatements due to the fact that an independent expert third party has performed periodic transaction and internal controls testing, and reviewed accounting choices made by management. Although the Call Report filings are not specifically audited by independent auditors, we make the assumption that a bank that is consistently audited should have higher quality financial reporting information than a bank that is consistently unaudited. However, we find that audited banks have a higher propensity to have restatements compared with peer unaudited banks. On a more positive note, we do find that the restatements of audited banks are of lower magnitude than unaudited banks.

Second, we use the estimated discretionary component of the allowance for loan losses as a proxy to measure the conservative recognition of loan losses and the magnitude of discretionary accruals used in reporting. We find that independent auditors do not promote financial reporting conservatism. Based on the study's estimation models, audited banks under-reported probable loan losses compared with unaudited banks. Consequently, audited banks were more aggressive in over-valuing assets and in over-reporting income as a result. This finding is counterintuitive as we expected audited banks to exhibit greater conservatism in recognizing potential losses. Auditors are more likely to be sued by being associated with over-valued assets, over-reported revenue, under-reported expenses, and under-reported losses in financial reporting (Pierre and Anderson 1984). We also find that audited banks used higher magnitudes of discretionary accruals in reporting. Higher magnitudes of discretionary accruals are of concern to auditors because they can cause estimation errors and therefore material misstatements (Francis and Krishnan 1999).

The remainder of the dissertation is organized as follows. In the next chapter, we discuss the literature and develop our hypotheses. In Chapter 3, we describe our sample and report descriptive statistics. Chapter 4 presents our research design and empirical models. Chapter 5 discusses our results and findings, and Chapter 6 elaborates on limitations, concludes the manuscript, and provides potential future research opportunities and directions.

Chapter 2

2. LITERATURE REVIEW & HYPOTHESES DEVELOPMENT

The objective of this chapter is to discuss the relevant literature and to develop our hypotheses. First, we review the use of machine learning techniques in the accounting and auditing literature. Second, we discuss the relationship between an independent audit and financial reporting quality.

2.1. Machine Learning in Accounting & Auditing

In the accounting and auditing literature, machine learning techniques are often used with varying level of success to predict or classify bankrupt, going concern, and fraudulent companies. In this section, we provide an overview of the specific machine learning techniques used in the literature for each of the three domains.

Bankruptcy is the inability for a company to pay back its debt obligations and commitments to creditors. In the US Bankruptcy Code, a company can reorganize under Chapter 11 where the company is considered a going concern or the company can liquidate under Chapter 7[4]. The literature has discussed numerous statistical, data mining, and machine learning techniques for predicting or forecasting the bankruptcy of companies. Some examples of these techniques range from uni-variate analysis (Beaver 1966), multiple discriminant analysis (Altman 1968), logit (Ohlson 1980) and probit

[4] Accounting Standards – 852 - Reorganizations

(Ohlson 1980; Zmijewski 1984) models, neural networks (Zhang et al. 1999; Tam and Kiang 1992; Odom and Sharda 1990), rough set theory (McKee 1998), discrete hazard models (Shumway 2001), instance based learners (Park and Han 2002), Bayesian models (Sarkar and Sriram 2001; Sun and Shenoy 2007; Aghaie and Saeedi 2009), rule learners (Thomaidis et al. 1999), decision trees algorithms (McKee and Greenstein 2000), genetic programming (McKee and Lensberg 2002), and support vector machines (Shin et al. 2005). In practice, the most commonly used technique is the discriminant analysis model called the ZETA credit risk model (Altman et al. 1977).

A going concern is issued by an auditor if the auditor has substantial doubt about the auditee's continual viability. A company is assumed a going concern if it is likely unable to continue meeting its future obligations as they become due (AU Section 341). In sequence, a going concern opinion should be issued by the auditor prior to the auditee filing for bankruptcy. However, in today's fast paced business environment this may not occur due to the annual nature of audits. If the auditor has substantial doubt about a company's ability to continue as a going concern, then the auditor expresses this opinion in an explanatory paragraph on the audit report. The statistical, data mining, and machine learning techniques used for predicting companies that are a going concern have been limited to a hand full. These techniques include logistic (Bell and Tabor 1991; Menon and Schwartz 1987; Mutchler 1985; Chen and Church 1992) and probit regression (Dopuch et al. 1987), matched sampling (Martens et al. 2008), and multiple discriminant analysis (Mutchler 1985; Levitan and Knoblett 1985). However, by far, the most popular going concern prediction technique used is logistic regression (Martens et al. 2008).

Management fraud is an intentional act to cause a material misstatement on the financial statements (AU Section 316). Fraud is more difficult to detect because management has no intention of leaving an obvious trail and there are often few examples of fraud to study (Kirkos et al. 2007). The literature commonly uses statistical, data mining, and machine learning techniques such as regression analysis (Abbott et al. 2000), logistic regression (Bell and Carcello 2000; Spathis et al. 2002; Beasley 1996), cascaded logit model (Summers and Sweeney 1998), generalized qualitative response model (Hansen et al. 1996), neural networks (Fanning and Cogger 1998; Green and Choi 1997; Kirkos et al. 2007), decision trees(Kirkos et al. 2007), bayesian belief networks (Kirkos et al. 2007), clustering (Thiprungsri 2011), and rule based models (Kim et al. 2011) to tease out management fraud. Recently, researchers and practitioners have been moving in the direction of using these techniques to continuously audit and monitor transactions within a company to detect fraud (Chan and Vasarhelyi 2011; Kim et al. 2011; Thiprungsri 2011). However, there is no dominating machine learning algorithm used for fraud detection.

Unlike predicting bankruptcy, going concern, and fraud, there is no practical need to predict or classify whether a bank is audited since the banks disclose this information voluntarily. However, we are interested to determine whether the decision to have an audit is systematic and endogenous. The accuracy of machine learning algorithms and using only bank characteristics to predict audited banks would provide evidence on whether the decision to have an audit is systematically determined and endogenous. To the best of our knowledge, there is no existing literature on the use of machine learning classification algorithms to predict or classify whether a private company is audited. Hence, we do not have any predetermined biased expectation on the ability or success of the use of machine learning algorithms in the classification of audited banks. However, (Kohlbeck 2005) does empirically model and analyze the demand for an audit. The statistical fit of their model was approximately 10%. They find that bank managers are more likely to choose an audit as the demand for monitoring and expertise increases. Hence, we slightly favor our expectation that machine learning algorithms should have some power in classifying audited and unaudited banks.

The discussions presented above can be summarized into the following hypothesis:

H1: *The decision to have an independent audit is systematic and endogenous.*

2.2. Auditing and Financial Reporting Quality

An independent audit can have a significant role in improving financial reporting quality (Cohen et al. 2004). However, an independent audit performed with low quality may be to the detriment of financial reporting quality. Hence, it can be argued that audit quality and financial reporting quality are positively correlated. Based on this assumption, the literature finds that audit and financial reporting quality is associated with material accuracy (DeAngelo 1981), conservative recognition of losses (Lee et al. 2006), and low magnitudes of discretionary accruals in reporting (Becker et al. 1998).

2.2.1. Material Accuracy and Restatements

Researchers and regulators have recognized that restatements by public companies are on the rise (Analytics 2007; Plumlee and Yohn 2009; Scholz 2008; Taub 2006; Turner and Weirich 2006). (Plumlee and Yohn 2009) find that during the period of 2003 to 2006, the primary cause of restatements was due to internal company errors and not the complexity of accounting standards. In the context of banks, (Lindo 2007) finds that the most common causes of restatements at public banking organizations were relating to the misapplication of accounting principles and calculation errors. The findings of (Plumlee and Yohn 2009) and (Lindo 2007) raises questions about the quality of audits being performed and the relationship between an independent audit and financial reporting quality. A restatement on audited financial statements is a symptom of the auditor's performance of a low quality audit. In other words, a restatement on issued audited financial statements implicates the auditor's failure to discover material misstatements during the audit. However, it is important to note that the inverse relationship does not hold true. If the audited financial statements are not restated, this does not indicate that a quality audit was performed. The company being audited could plausibly have had excellent controls over financial reporting irrespective of a low quality audit or that a material misstatement exists but was not discovered subsequently.

A restatement can also be argued to be a negative joint reflection on the effectiveness of financial reporting controls maintained by management and the effectiveness of the audits performed by external auditors. As part of SOX 2002, management and their independent auditors of public companies are required to assess the adequacy of the internal controls over financial reporting[5]. Although a joint responsibility exists, ultimately an independent audit is supposed to serve as a compensating control to mitigate the ineffectiveness of management's internal controls over financial reporting and thus the flow of material inaccuracies into financial statements. Moreover, the auditors have a responsibility to perform an audit with professional due care. A quality audit was not performed if it failed to detect and correct material misstatements during the audit (DeAngelo 1981). (Turner and Sennetti 2001) find that the probability of a restatement is decreased when a quality audit is performed. Hence, a company that is consistently audited should have a higher probability that their financial results are more reliable than an unaudited company's financial information. Furthermore, the magnitude of those misstatements on audited financial statements is expected to be lower than unaudited financial statements.

The discussions presented above can be summarized into the following hypotheses:

[5] Sarbanes Oxley Act 2002 – Section 404 (Assessment of Internal Controls)

H2: *An independent audit decreases the likelihood of having restatements.*

H3: *An independent audit decreases the magnitudes of restatements.*

2.2.2. Conservatism and Magnitude of Discretionary Accruals

For banks, the allowance for loan losses is the largest accounting estimate. The literature recognizes that the allowance for loan losses account consist of a nondiscretionary and a discretionary component (Beaver and Engel 1996). These two components are not directly observable to external users of financial statements. The discretionary component is comprised of management's subjective judgment on the quality of loans. Using discretion, management may deviate from conservative estimation of loan portfolio quality by intentionally under reporting (aggressive) probable loan losses. Moreover, bank managers may use the estimation of loan portfolio quality as a tool to manage earnings (Beaver and Engel 1996; Beatty et al. 2002). The un-conservative estimation of loan portfolio quality can cause misleading financial results. (Dahl et al. 1998) find that auditors are particularly concerned with the valuation of the allowance for loan losses on their effect on net book value and income. As a gatekeeper, the auditor is responsible for assessing the allowance for loan losses and the loan loss provision account for material accuracy, compliance with GAAP, excessiveness, presentation of losses, and adequate disclosure (AICPA 2004).

Under Generally Accepted Accounting Principles (GAAP), an impairment of loans should be recognized when it is probable that a credit loss has been incurred[6]. According to (Basu 1997), conservatism in financial reporting is the recognition of losses more quickly than gains. The timely recognition of losses can have a significant impact on earnings quality (Ball and Shivakumar 2005). Under conservative reporting, management should discretionarily bias towards over accruing for probable loan losses to prevent the potential underreporting of losses and the over reporting of income. Furthermore, if a company is audited, we can expect more conservative financial reporting or the discretionary over accruing for potential loan losses. A quality audit is associated with conservatism in financial reporting (Lee et al. 2006). Auditors have a tendency to promote the conservative recognition of gains and losses (Cano-Rodríguez 2010; Basu et al. 2002; Chung et al. 2003; Becker et al. 1998). Legal liability plays a critical role in encouraging auditors to be more conservative in their assessments (Francis and Wang 2004; Francis et al. 2003; DeFond and Subramanyam 1998). The likelihood of litigation increases when the audit client under accruals for potential losses (Heninger 2001) or when there are material misstatements that result in over-valuing assets, over-reporting revenue, and under-reporting expenses and liabilities (Pierre and Anderson 1984). Furthermore, a more conservative audit is expected to be performed by auditors when there is regulatory oversight (Van Tendeloo and Vanstraelen 2008).

After the enactment of the Sarbanes Oxley Act 2002 (SOX), (Lobo and Zhou 2006; Zhou 2007) finds public companies became more conservative in the recognition of

[6] FASB ASC 310-10-35-4

losses and began reporting lower magnitudes of discretionary accruals. A potential reason for this effect could be attributed to SOX increasing financial reporting legal ramifications and repercussions on management and their auditors. Generally Accepted Accounting Principles (GAAP) allows management to report a range of acceptable values for accounting estimates because of their subjective nature. However, if irregular or erroneous excess discretion is used by management then current and future earnings quality can be eroded. Hence, the magnitude of discretion used in financial reporting should be limited. For those companies who are audited, we can expect the more conservative use of discretionary accruals. For auditors, audit risk escalates when higher magnitudes of discretion is used in reporting because material estimation errors can occur (Francis and Krishnan 1999). Furthermore, (Bartov et al. 2000; Becker et al. 1998) find that clients of Big Six auditors had lower magnitudes of discretionary accruals than non-Big Six clients. The Big Four audit firms are known to be synonymous with audit quality. Thus, it can be inferred that a company that is consistently audited with quality should be associated with lower magnitudes of discretionary accruals in financial reporting.

The discussions presented above can be summarized into the following hypotheses:

H4: *An independent audit increases the conservative recognition of expected and probable losses.*

H5: *An independent audit decreases the magnitudes of discretionary accruals in financial reporting.*

Chapter 3

3. SAMPLE AND DESCRIPTIVE STATISTICS

The data used in this study is obtained from the Report of Condition and Income (Call Reports). US commercial banks are required to submit a Call Report to the FDIC on a quarterly basis. The Call Reports are used by regulators to monitor the performance and stability of banks. In our study, we use the fourth quarter Call Report data (annualized)[7]. The Call Report is due from the commercial bank within 30 days after the end of each calendar quarter[8]. The financial reporting requirements imposed by the FDIC for commercial banks are similar to those required by the Securities and Exchange Commission (SEC) for public companies. The regulatory reports are based on Generally Accepted Accounting Principles (GAAP) rather than on statutory accounting principles (Beatty and Harris 1999) and are examined on a regular basis by regulators (Gunther and Moore 2003).

An advantage of analyzing small private commercial banks utilizing Call Report data is twofold. First, we have a large dataset which consist of audited and unaudited banks due to no regulatory requirement of an independent audit. Small private commercial banks with less than $500 million in total assets are exempt from the

[7] FFIEC 031 and 041
[8] FFIEC 031 and 041

mandate of having an annual independent audit[9,10]. The option to have an audit by a homogenous large population allows us to analyze in isolation the impact that an independent audit has on financial reporting quality. Second, smaller companies have a tendency to have lower quality financial reporting and therefore increase the benefit of having an independent audit. The literature finds that smaller companies within industries are specifically found more likely to have restatements and hence have lower financial reporting quality (Kinney and McDaniel 1989; Plumlee and Yohn 2010; Turner and Sennetti 2001).

The largest bank during the period under study has approximately $499,928,000 in total assets and hence all the banks in this study are exempted by regulation from having an independent audit. We only use banks with under $500 million in total assets for all 10 years under study. The data used ranges from 2001 to 2010 (10 Years) and includes only private US commercial banks[11]. The audit indicator variable provided by the Call Reports consists of eight categories representing the highest level of accounting or auditing service obtained[12]. We construct our binary independent audit variable using the first two categories. A commercial bank is categorized as audited if the commercial bank (Category 1) or its parent bank holding company (Category 2) had an

[9] FDI Act 12 C.F.R. Part 363

[10] 3 Exceptions: 1) Thrifts, regardless of size, with a composite safety and soundness CAMELS rating of 3,4,5; 2) Holding companies which control insured financial institution subsidiary(ies) with aggregate consolidated assets of $500 million or more; 3) Any other entity for which the OTS determines an audit is required for safety and soundness reasons; (OTS Audit Rule 12 CFR 562.4) and 4) Small banks part of a public holding company.

[11] The years 2001 and 2010 were removed in analysis because of need for lagging and leading of variables for certain calculations and variable transformations.

[12] FFIEC 031 and 041

independent audit. There are 3,962 unique commercial banks in the dataset and 2,233 of them are audited. Therefore, approximately 56% of the banks decided to have an audit absent regulatory requirements. All banks were active and their decision to have an audit or to not have an audit was persistent for all 10 years. The purpose of using persistent banks is because the Call Reports are not directly audited but the financial information used to generate those Call Reports is audited. Hence, we make the assumption that a bank that consistently obtains an independent audit should have higher quality financial information than a consistently unaudited bank. **Table 1A/B/C/D** provides summary statistics and **Table 2** provides Pearson's correlation coefficients for all the variables used in this study. All the variables used in this study were standardized during analysis to provide comparable and interpretable coefficients.

We also construct our main dependent variables using the restatement and the allowance for loan losses variable from the Call Report. The FDIC defines a restatement as corrections resulting from; (1) mathematical mistakes, (2) mistakes in applying accounting principles, (3) improper use of information which existed when the prior Reports of Condition and Income were prepared, and (4) a change from an accounting principle that is neither accepted nor sanctioned by bank supervisors to one that is acceptable to supervisors. We transform the continuous numeric restatement variable into two constructed variables; the occurrence of a restatement (dummy variable) and the magnitude of a restatement (absolute value of restatement variable). **Table 3** shows numerically and **Figure 1** shows graphically the trend in number of restatements over the period of study. Notably there was a spike in number of restatements from 2007 to 2009. We cannot find an explanation for the spike in number of restatements. However, we cannot speculate that it was due to estimation errors of loan portfolio quality from the mortgage crisis since restatements do not specifically include estimation errors. **Table 4**, **Table 5**, and **Figure 2** show the breakdown of restatements between unaudited and audited banks. The univariate analysis clearly shows that audited banks had more restatements than unaudited banks. And **Table 6** shows the number of banks that have multiple restatements over the period under study. Interestingly, there are fifty two banks that have five or more restatements over a 10 year span (34 audited and 18 unaudited).

The FDIC defines the allowance for loan losses variable to be a reserve amount for possible loan losses. The literature recognizes the allowance for loan losses account consist of two components. Using the (Beaver and Engel 1996) estimation model, we estimate the non-discretionary and discretionary component of the allowance for loan losses account. We then transform the continuous numeric discretionary component of the allowance for loan losses variable into two constructed variables; discretionary conservative recognition of probable loan losses (dummy variable - sign of residual from the (Beaver and Engel 1996) model), and magnitude of discretionary accrual (absolute value of the residual from the (Beaver and Engel 1996) model).

Figure 3 and **Figure 4** show the number of audited and unaudited banks who were not conservative in recognizing loan losses and the magnitude of discretionary accruals over the period under study, respectively. Un-conservative recognition of probable loan losses is defined as a negative value for the residual created by the (Beaver and Engel 1996) model. We can see from **Figure 3** that audited banks were consistently less conservative in recognizing potential loan losses than unaudited banks. However, from **Figure 4** we see inconsistency in the magnitude of discretionary accruals used between audited and unaudited banks. But noticeably, audited banks had higher magnitudes of discretionary accruals towards the end of the study period. We examine these univariate findings further using multivariate statistical models in the next chapter.

Table 1

Summary Statistics

A. Unaudited Banks

Variable	N	Mean	Std Dev	Median	Minimum	Maximum
LTA	13548	-0.48533	0.93744	0.47569	-4.35524	2.00091
OFF	13548	2.79569	2.67816	2	1	65
NIITI	13548	0.10013	0.05952	0.08915	-0.28164	0.90997
OBTA	13548	0.08566	0.06837	0.07249	0	2.43242
MU	13548	0.000664	0.02577	0	0	1
BHC	13548	0.84765	0.35937	1	0	1
GR	13548	0.05696	0.13088	0.04132	-0.41402	3.60613
ROAA	13548	0.0108	0.00812	0.01068	-0.10754	0.15583
LTC	13548	6.0647	2.45162	6.05553	0.08918	31.47715
DALL	13548	-0.04196	0.90335	0.13158	-1.89258	9.07786
DCO	13548	-0.04096	0.9087	0.28943	-0.46279	36.85237
DTL	13548	-0.10804	0.98916	0.09914	-2.63317	6.25032
DNPL	13548	-0.06532	0.80651	-0.3373	-0.58571	13.29604
ΔDNPL	13548	-0.07024	0.85132	0.14746	-15.5803	9.92234
SLTL	13548	5.73229	2.30181	5.72657	0.09204	34.6
SΔNPL	13548	0.00803	0.08066	0	-1.4197	1.21253
SLNPL	13548	0.05271	0.09038	0.02476	0	1.90811
SLALL	13548	0.07996	0.04359	0.07473	0.00134	1.10909
AI	13548	0	0	0	0	0
RS	13548	0.05669	0.23125	0	0	1
rALLDirect	13548	0.45675	0.49814	0	0	1
rALLABS	13548	0.5256	0.4961	0.39972	6.01E-06	7.75156

B. Audited Banks

Variable	N	Mean	Std Dev	Median	Minimum	Maximum
LTA	17602	0.37361	0.87889	0.47772	-4.1375	2.006
OFF	17602	4.41052	3.12779	4	0	29
NIITI	17602	0.12267	0.1059	0.10582	-2.30321	0.99937
OBTA	17602	0.75605	16.13589	0.08485	0	785.3808
MU	17602	0.06494	0.24642	0	0	1
BHC	17602	0.80826	0.39368	1	0	1
GR	17602	0.07809	0.2956	0.05001	-0.89425	27.8154
ROAA	17602	0.00955	0.0192	0.00955	-0.52991	0.71781
LTC	17602	6.56649	3.85219	6.59105	0	391.4762
DALL	17602	0.03199	1.06587	-0.08698	-1.94445	21.18037
DCO	17602	0.0289	1.00422	-0.26026	-0.46279	30.81058
DTL	17602	0.08326	1.0002	0.12056	-2.67407	18.76544
DNPL	17602	0.05034	1.12399	-0.29361	-0.58571	30.47707
ΔDNPL	17602	0.05404	1.09786	-0.13531	-13.669	18.75628
SLTL	17602	6.12628	4.67347	6.12919	0	540.5329
SΔNPL	17602	0.01988	0.26933	0.000751	-1.31225	33.01587
SLNPL	17602	0.05836	0.1287	0.027	0	9.37188
SLALL	17602	0.08167	0.07725	0.07431	0	6.82313
AI	17602	1	0	1	1	1
RS	17602	0.07306	0.26024	0	0	1
rALLDirect	17602	0.42336	0.49411	0	0	1
rALLABS	17602	0.53312	0.56471	0.3965	6.59E-05	10.4389

Note: Definitions of the variables is available on the next page and in the **APPENDIX**.

C. Unaudited Banks with Restatements

Variable	N	Mean	Std Dev	Median	Minimum	Maximum
RSABS	768	92.60286	357.685	24	1	7498

D. Audited Banks with Restatements

Variable	N	Mean	Std Dev	Median	Minimum	Maximum
RSABS	1286	144.5451	361.4735	52	1	9630

Variable Definitions

LTA	=	Log of Total Assets;
OFF	=	Number of offices, branches, locations, and facilities;
NIITI	=	Non-interest Income divided by total Interest and Non-interest income;
OBTA	=	Off-balance sheet activities divided by total assets;
MU	=	Mutual or stockholder bank (Dummy Variable - 1 if a mutual bank and 0 a stock bank);
BHC	=	Parent is a Bank Holding Company (Dummy Variable - 1 if the bank's parent is a bank holding company and 0 otherwise);
GR	=	Change in total assets divided by beginning total assets;
ROAA	=	Net income (Loss) divided by Total Average Assets (assets at the end of the previous year plus assets at the end of the current year divided by 2);
LTC	=	Total loans divided by Total Equity Capital;
DALL	=	Allowance for loan losses;
DCO	=	Loan charge offs;
DTL	=	Total Loans;
DNPL	=	Nonperforming loans;
ΔDNPL	=	Change in nonperforming loans as a percentage of the average of beginning and ending total loans;
SLTL	=	Beginning balance of total loans;
SΔNPL	=	Change in Nonperforming Loans;

SLNPL	=	Beginning balance of nonperforming loans;
SLALL	=	Beginning balance of allowance for loan losses;
AI	=	Audit Indicator (Dummy Variable - 1 if the bank is independently audited and 0 otherwise);
RS	=	Restatement (Dummy Variable - 1 if the bank restated and 0 otherwise);
RSABS	=	Absolute value of Restatement;
rALLDirect	=	Discretionary component of allowance for loan losses (Residual from (Beaver and Engel 1996) Model);
rALLABS	=	Magnitude of discretionary component of allowance for loan losses (Absolute value of residual from (Beaver and Engel 1996) Model;

Table 2

Pearson Correlation Table

	LTA	OFF	NIITI	OBTA	MU	BHC	GR	ROAA	LTC	DALL	DCO	DTL	DNPL	ΔDNPL	SLTL	SΔNPL	SLNPL	SLALL	AI	RS	RSABS	rALLDirect	rALLABS
LTA	1	0.59153	0.01587	-0.05754	0.06333	0.15786	0.05889	-0.02304	0.22821	0.11554	0.06763	0.30596	0.11483	0.1033	0.17168	0.05256	0.04166	0.03771	0.42582	0.0346	0.05256	-0.04294	-0.05146
		<.0001	0.0051	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
OFF	0.59153	1	0.15824	-0.02602	-0.03482	0.15166	0.0511	-0.01967	0.16088	0.09299	0.04712	0.22126	0.05671	0.0384	0.1187	0.01439	0.03176	0.03973	0.26268	0.03278	0.02487	-0.00706	-0.05911
	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	0.0005	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0111	<.0001	<.0001	<.0001	<.0001	<.0001	0.2126	<.0001
NIITI	0.01587	0.15824	1	0.28491	-0.0975	0.02423	0.00324	0.28334	-0.06229	-0.04837	-0.00277	-0.08988	-0.03504	-0.03198	-0.04767	-0.00529	-0.01824	-0.01226	0.12491	0.00959	0.00503	-0.01911	-0.00752
	0.0051	<.0001		<.0001	<.0001	<.0001	0.5677	<.0001	<.0001	<.0001	0.6253	<.0001	<.0001	<.0001	<.0001	0.3501	0.0013	0.0305	<.0001	0.0904	0.3749	0.0007	0.1846
OBTA	-0.05754	-0.02602	0.28491	1	-0.00636	-0.06488	-0.00361	0.351	-0.05309	-0.04446	-0.00229	-0.07453	-0.01622	-0.00406	-0.04301	-0.00221	-0.01317	-0.02705	0.02739	0.01998	0.01031	-0.017	0.01284
	<.0001	<.0001	<.0001		0.2613	<.0001	0.5239	<.0001	<.0001	<.0001	0.6859	<.0001	0.0042	0.4741	<.0001	0.6961	0.0201	<.0001	<.0001	0.0004	0.0688	0.0027	0.0235
MU	0.06333	-0.03482	-0.0975	-0.00636	1	-0.41173	-0.0212	-0.06761	-0.02736	-0.16293	-0.065	-0.01848	-0.02223	0.00021	-0.01785	-0.00037	-0.02958	-0.11119	0.16884	0.03635	0.0318	-0.13661	0.08619
	<.0001	<.0001	<.0001	0.2613		<.0001	0.0002	<.0001	<.0001	<.0001	<.0001	0.0011	<.0001	0.9701	0.0016	0.9478	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
BHC	0.15786	0.15166	0.02423	-0.06488	-0.41173	1	0.02451	0.00148	0.15535	0.17052	0.04229	0.19972	0.01812	0.00692	0.1222	0.00904	0.01167	0.10633	-0.05144	-0.03147	-0.00314	0.09045	-0.05822
	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	0.7942	<.0001	<.0001	<.0001	<.0001	0.0014	0.222	<.0001	0.1107	0.0395	<.0001	<.0001	<.0001	0.5797	<.0001	<.0001
GR	0.05889	0.0511	0.00324	-0.00361	-0.0212	0.02451	1	-0.00726	0.0524	-0.01853	-0.07128	0.08375	-0.06004	0.04286	-0.04226	-0.00338	-0.07867	-0.07172	0.04391	-0.00422	-0.00029	-0.007	-0.04044
	<.0001	<.0001	0.5677	0.5239	0.0002	<.0001		0.2002	<.0001	0.0011	<.0001	<.0001	<.0001	<.0001	<.0001	0.5505	<.0001	<.0001	<.0001	0.4565	0.9598	0.2165	<.0001
ROAA	-0.02304	-0.01967	0.28334	0.351	-0.06761	0.00148	-0.00726	1	-0.11098	-0.20987	-0.29966	-0.10404	-0.24842	0.00903	-0.11539	-0.07088	-0.21188	-0.1267	-0.0403	-0.03475	-0.01022	-0.01127	-0.1276
	<.0001	0.0005	<.0001	<.0001	<.0001	0.7942	0.2002		<.0001	<.0001	<.0001	<.0001	<.0001	0.111	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0713	0.0468	<.0001
LTC	0.22821	0.16088	-0.06229	-0.05309	-0.02736	0.15535	0.0524	-0.11098	1	0.51335	0.23901	0.78001	0.29742	0.05555	0.95321	0.63373	0.45667	0.65513	0.0748	-0.00384	0.00012	0.10302	0.1283
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.4981	0.983	<.0001	<.0001
DALL	0.11554	0.09299	-0.04837	-0.04446	-0.16293	0.17052	-0.01853	-0.20987	0.51335	1	0.47431	0.4801	0.45868	-0.00736	0.47256	0.19883	0.44912	0.71037	0.03669	0.00247	0.02419	0.56394	0.45029
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0011	<.0001	<.0001		<.0001	<.0001	<.0001	0.194	<.0001	<.0001	<.0001	<.0001	<.0001	0.6624	<.0001	<.0001	<.0001
DCO	0.06763	0.04712	-0.00277	-0.00229	-0.065	0.04229	-0.07128	-0.29966	0.23901	0.47431	1	0.16572	0.48701	-0.04291	0.28698	0.15362	0.49806	0.47955	0.03591	0.01827	0.03157	0.0075	0.31872
	<.0001	<.0001	0.6253	0.6859	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0013	<.0001	0.1858	<.0001
DTL	0.30596	0.22126	-0.08988	-0.07453	-0.01848	0.19972	0.08375	-0.10404	0.78001	0.4801	0.16572	1	0.21216	0.10984	0.60217	0.09796	0.1645	0.31183	0.09485	-0.00522	-0.00479	0.07303	0.05591
	<.0001	<.0001	<.0001	<.0001	0.0011	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.3569	0.3977	<.0001	<.0001
DNPL	0.11483	0.05671	-0.03504	-0.01622	-0.02223	0.01812	-0.06004	-0.24842	0.29742	0.45868	0.48701	0.21216	1	-0.17606	0.33203	0.4242	0.65595	0.36805	0.05734	0.04113	0.02288	-0.00257	0.26693
	<.0001	<.0001	<.0001	0.0042	<.0001	0.0014	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.6503	<.0001
ΔDNPL	0.1033	0.0384	-0.03198	-0.00406	0.00021	0.00692	0.04286	0.00903	0.05555	-0.00736	-0.04291	0.10984	-0.17606	1	0.01265	-0.08358	-0.11238	-0.04124	0.06161	0.00777	0.00049	0.01462	-0.03276
	<.0001	<.0001	<.0001	0.4741	0.9701	0.222	<.0001	0.111	<.0001	0.194	<.0001	<.0001	<.0001		0.0256	<.0001	<.0001	<.0001	<.0001	0.1704	0.9312	0.0099	<.0001
SLTL	0.17168	0.1187	-0.04767	-0.04301	-0.01785	0.1222	-0.04226	-0.11539	0.95321	0.47256	0.28698	0.60217	0.33203	0.01265	1	0.74776	0.54847	0.72706	0.05097	-0.00387	0.00192	0.08935	0.14798
	<.0001	<.0001	<.0001	<.0001	0.0016	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0256		<.0001	<.0001	<.0001	<.0001	0.4944	0.7353	<.0001	<.0001
SΔNPL	0.05256	0.01439	-0.00529	-0.00221	-0.00037	0.00904	-0.00338	-0.07088	0.63373	0.19883	0.15362	0.09796	0.4242	-0.08358	0.74776	1	0.37465	0.54115	0.02805	0.01033	0.00573	-0.0072	0.09729
	<.0001	0.0111	0.3501	0.6961	0.9478	0.1107	0.5505	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	0.0684	0.3116	0.2038	<.0001
SLNPL	0.04166	0.03176	-0.01824	-0.01317	-0.02958	0.01167	-0.07867	-0.21188	0.45667	0.44912	0.49806	0.1645	0.65595	-0.11238	0.54847	0.37465	1	0.61353	0.02463	0.02198	0.01558	0.04711	0.28381
	<.0001	<.0001	0.0013	0.0201	<.0001	0.0395	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.0003	0.006	<.0001	<.0001
SLALL	0.03771	0.03973	-0.01226	-0.02705	-0.11119	0.10633	-0.07172	-0.1267	0.65513	0.71037	0.47955	0.31183	0.36805	-0.04124	0.72706	0.54115	0.61353	1	0.01303	-0.00144	0.00741	0.34319	0.36054
	<.0001	<.0001	0.0305	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		0.0215	0.7988	0.1912	<.0001	<.0001
AI	0.42582	0.26268	0.12491	0.02739	0.16884	-0.05144	0.04391	-0.0403	0.0748	0.03669	0.03591	0.09485	0.05734	0.06161	0.05097	0.02805	0.02463	0.01303	1	0.03271	0.02773	-0.03336	0.00696
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0215		<.0001	<.0001	<.0001	0.2196
RS	0.0346	0.03278	0.00959	0.01998	0.03635	-0.03147	-0.00422	-0.03475	-0.00384	0.00247	0.01827	-0.00522	0.04113	0.00777	-0.00387	0.01033	0.02198	-0.00144	0.03271	1	0.03059	-0.0194	0.03443
	<.0001	<.0001	0.0904	0.0004	<.0001	<.0001	0.4565	<.0001	0.4981	0.6624	0.0013	0.3569	<.0001	0.1704	0.4944	0.0684	0.0003	0.7988	<.0001		<.0001	0.0006	<.0001
RSABS	0.05256	0.02487	0.00503	0.01031	0.0318	-0.00314	-0.00029	-0.01022	0.00012	0.02419	0.03157	-0.00479	0.02288	0.00049	0.00192	0.00573	0.01558	0.00741	0.02773	0.03059	1	-0.00619	0.03947
	<.0001	<.0001	0.3749	0.0688	<.0001	0.5797	0.9598	0.0713	0.983	<.0001	<.0001	0.3977	<.0001	0.9312	0.7353	0.3116	0.006	0.1912	<.0001	<.0001		0.275	<.0001
rALLDirect	-0.04294	-0.00706	-0.01911	-0.017	-0.13661	0.09045	-0.007	-0.01127	0.10302	0.56394	0.0075	0.07303	-0.00257	0.01462	0.08935	-0.0072	0.04711	0.34319	-0.03336	-0.0194	-0.00619	1	0.12422
	<.0001	0.2126	0.0007	0.0027	<.0001	<.0001	0.2165	0.0468	<.0001	<.0001	0.1858	<.0001	0.6503	0.0099	<.0001	0.2038	<.0001	<.0001	<.0001	0.0006	0.275		<.0001
rALLABS	-0.05146	-0.05911	-0.00752	0.01284	0.08619	-0.05822	-0.04044	-0.1276	0.1283	0.45029	0.31872	0.05591	0.26693	-0.03276	0.14798	0.09729	0.28381	0.36054	0.00696	0.03443	0.03947	0.12422	1
	<.0001	<.0001	0.1846	0.0235	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.2196	<.0001	<.0001	<.0001	

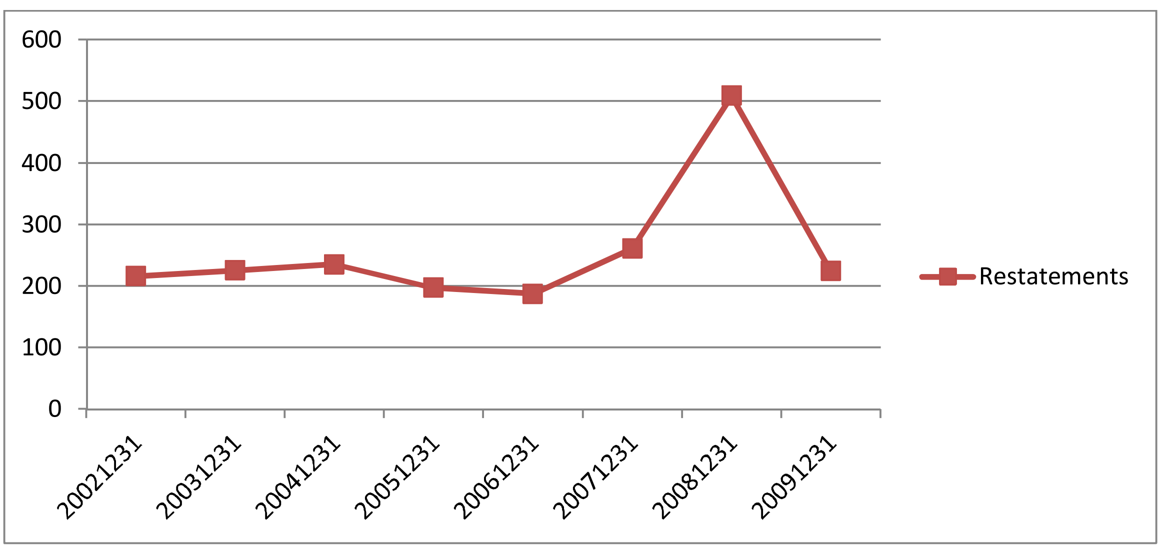
Table 3

Number of Restatements by Year

Year	Restatements
2002	216
2003	225
2004	235
2005	197
2006	187
2007	261
2008	509
2009	224
Total	2054

Figure 1

Number of Restatements by Year



X-Axis = Date

Y- Axis = Number of Restatements

Table 4

Number of Restatements by Unaudited Banks

Year	Restatements
2002	104
2003	105
2004	95
2005	90
2006	69
2007	80
2008	140
2009	85
Total	768

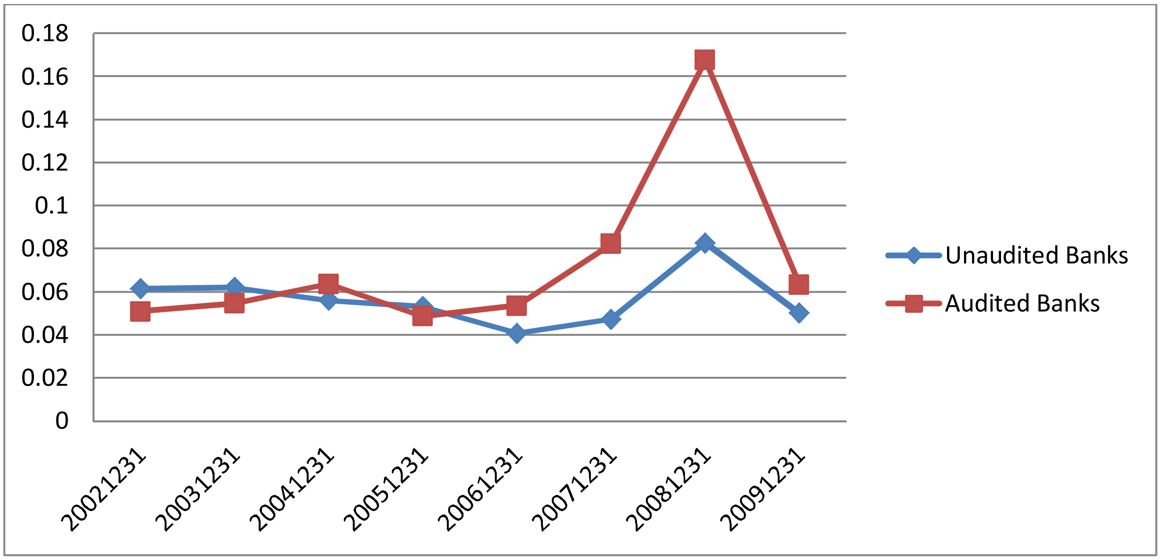
Table 5

Number of Restatements by Audited Banks

Year	Restatements
20021231	112
20031231	120
20041231	140
20051231	107
20061231	118
20071231	181
20081231	369
20091231	139
Total	1286

Figure 2

Frequency of Restatements (Audited vs. Unaudited)



X-Axis = Date

Y- Axis = Number of Restatements Over Total Restatements

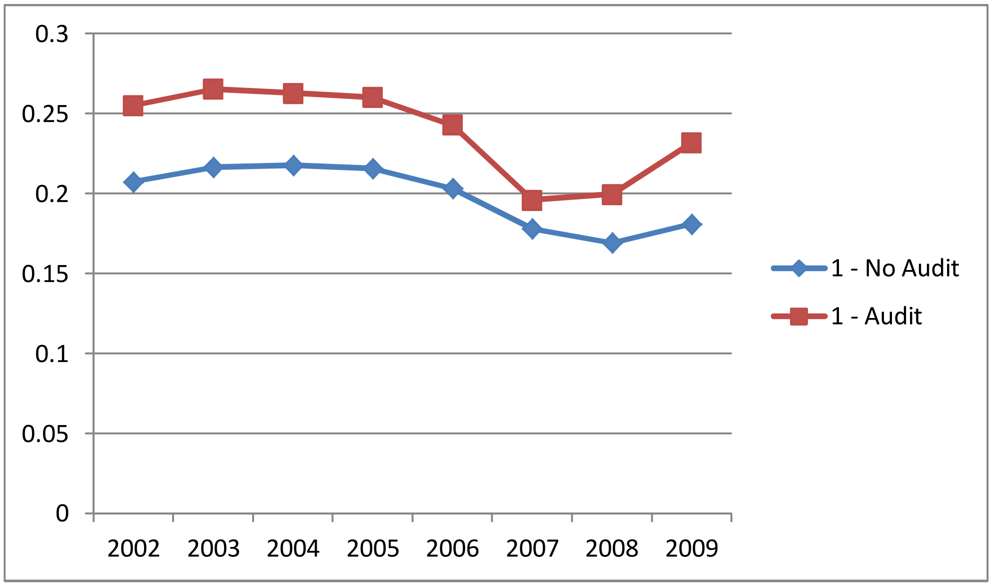
Table 6

Number of Banks with Multiple Restatements (Unaudited vs. Audited)

# of Restatements	1	2	3	4	5	6	7	8	Total
Not Audited - 0	314	98	33	14	10	4	3	1	477
Audited - 1	511	155	55	23	13	8	9	4	778
Total	825	253	88	37	23	12	12	5	1255

Figure 3

Frequency of Un-conservative Recognition of Losses

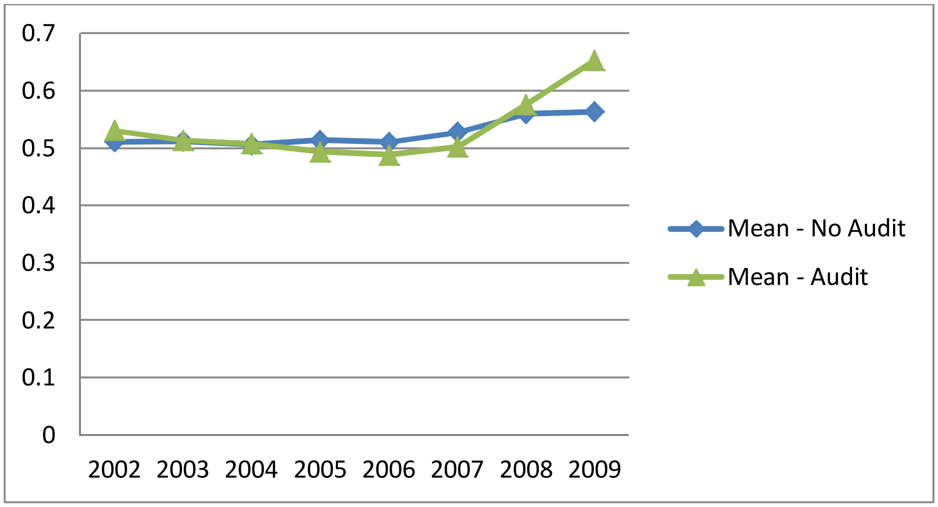


X-Axis = Date

Y- Axis = Number of Banks with Un-conservative Recognition of Losses Over Total Banks

Figure 4

Mean of Magnitude of Discretionary Accruals



X-Axis = Date

Y- Axis = Mean of Residual (Discretionary Accrual) from (Beaver and Engel 1996) model (using standardized values).

Chapter 4

4. RESEARCH DESIGN AND EMPIRICAL MODELS

4.1.Machine Learning Algorithms

Our data set consists of labeled audited and unaudited banks and hence can be viewed and used for a supervised classification problem. The classification process consists of training a machine learning classification model with training data and applying the model to unseen test data to evaluate prediction accuracy (Tan et al. 2005). The number of observations (bank years) that have an audit (17,602) or did not have an audit (13,548) is relatively balanced in our dataset and works in our favor **(Table 7)**. An imbalance class may cause the accuracy measures of the models to be suspect. SAS[13] and Weka software (Witten and Frank 2005) were used to manipulate and analyze the data, respectively. Using the methodology of 10-fold cross-validation, Weka's Experimenter program was used to test and compare the performance of six learning algorithms on predicting or classifying whether a bank had an independent audit. A 10-fold cross-validation was used to identify potential over-fitting in the model and to reduce error variance. In a 10-fold cross-validation, we divide a dataset into 10 random subsets. A subset is set aside for testing and each learning algorithm is trained with the other 9 subsets. The trained model is then tested on the omitted testing subset. This process is repeated 10 times, each time setting aside a different testing subset, and then the results are averaged.

[13] http://www.sas.com

Table 7

Audit Indicator

Audit Status	Number Unique Banks	Observations
Audited	2,233	17,602
Unaudited	1,729	13,548
Total	3,962	31,150

The six learning algorithms used in this study include: Naïve Bayes (John and Langley 1995), Simple Logistic (Sumner et al. 2005), Support Vector Machine (Hastie and Tibshirani 1998; Keerthi et al. 2001; Platt 1999), IBK (Aha et al. 1991), JRIP (Cohen et al. 2002), and Random Forest (Breiman 2001). In making classifications, the Naïve Bayes algorithm considers each predictor variable used in the model independently and aggregates the probability of each variable's contribution to being in a class. In the Support Vector Machine algorithm, a linear hyperplane is used to separate the classes. A perfect linear separation occurs when the different class test examples lie on different sides of the hyperplane with the greatest margin from the hyperplane. The IBK algorithm is a k-nearest neighbor's algorithm (k-NN). The k-nearest neighbor algorithm is a method for classifying test examples based on the closest training example(s) in a dimensional space. A test example's proximity is compared with other training examples using Euclidean Distance (jaccard and cosine similarity measures can also be used). The test example is then classified as the majority class of the k-nearest neighbors. The JRIP algorithm (aka Ripper) is a rule based classifier that uses a class based ordering scheme. JRIP extracts a set of rules from the training set that identifies the relationship between predictors and the predicted classes. Lastly, the Random Forest learning algorithm is an ensemble classifier. This learning algorithm builds multiple decision trees using random vectors and then uses the majority vote of the decision trees to classify the test examples.

In training and testing our six machine learning algorithms, we use characteristics denoted by the literature that may increase the likelihood that a company may have an independent audit. The literature cites size, hierarchical and ownership structure, growth, complex operations, and profitability as key characteristics that may influence a company's decision to have a voluntary audit. As a company grows in size it becomes more difficult for owners and managers to be observant of all facets of operations. Hence, (Tauringana and Clarke 2000; Chow 1982; Abdel-Khalik 1993) find that as the company size increases the likelihood of procuring an audit increases. The log of total assets is used to proxy for size (LTA). The hierarchical structure of a company (Abdel-Khalik 1993) and managerial ownership (Tauringana and Clarke 2000) can also influence the decision to have an independent audit. Hierarchical structure is represented by whether the commercial bank is part of a bank holding company (BHC) and managerial ownership is represented by whether the bank is owned by depositors or stockholders (MU).

As a company grows they may need to seek out financial experts because of the lack of in house financial expertise (Aier et al. 2005). We use change in total assets to proxy for growth (GR). Furthermore, the complexity of a company's operations can also affect the decision to have an audit (Kohlbeck 2005). External auditors can provide expertise to guide the accounting for complex operations because they generally have industry specific expertise and the experience of auditing peer companies. Non-interest income over total income (NIITI) (Kohlbeck 2005) and off balance sheet activity (OBTA) are used as proxies for the complexity of banking operations. Banks are traditionally in the business of making interest income and the accounting complexity for non-interest earning activities may be more challenging. Furthermore, off balance sheet activity can be complicating to account for and may warrant the decision to procure a financial accounting expert. Lastly, (Kreutzfeldt and Wallace 1987) find that companies with profitability problems have larger and more frequent accounting errors and hence may seek out an audit to have more reliable financial information. Return on average assets (ROAA) is used to represent profitability.

4.2.Restatements

We use Probit (**Model 1**) and Gamma Regression (**Model 2**) to model the relationship between procuring an independent audit with the likelihood of having a restatement and the magnitude of those restatements, respectively. A Probit model is used because the restatement dependent variable is a dummy variable/binary (Imai et al. 2007b) and a Gamma Model is used because the dependent variable magnitude of restatement is positively distributed (Imai et al. 2007a). Common accounting and banking related control variables that may increase or decrease the likelihood of a restatement are used in the model. The log of total assets (LTA) is used to control for the bank's size effect on the likelihood of having a restatement. We expect that larger banks have better internal controls and management with greater financial expertise than smaller banks. The number of offices (OFF) (Kohlbeck 2005), non-interest income over total income (NIITI) (Kohlbeck 2005) and off balance sheet activity (OBA) are used as proxies for the complexity of banking operations. The complexity of banking operations may increase the likelihood of having a restatement. A greater number of branches and offices can impose greater complexity in operations and accounting. Likewise, since commercial banks traditionally earn their income through interest products and services, banks involved in non-interest activities such as issuing insurance and brokering securities can introduce more complex accounting issues and lead to a higher likelihood of misstatements. Additionally, banks with off balance sheet activities can be more difficult to account for and classify and thus may increase the likelihood of restatements.

We use the variables (MU) and (BHC) as a proxy for ownership and hierarchical structure type, respectively. (MU) indicates whether the bank is owned by its depositors (Mutual Bank) or stockholders (Stockholder Bank). Ownership and hierarchical type may influence corporate governance over financial reporting and hence decrease the likelihood of a restatement. We assume that depositors have very little interest in managing the day to day activities of a bank or even the performance of the banks since their deposits are federally insured. Hence, depositors will not have a demand for quality financial reporting. However, since we are studying small private commercial banks, there is a higher likelihood that the owners of stockholder banks are engaged in the day to day activities of the bank. Hence, owners of stockholder banks may have a greater interest in higher quality financial information for measuring the performance of the bank. The (BHC) variable indicates whether the commercial bank is part of a bank holding company. Unlike stockholder banks, management of the bank holding company is likely removed from the day to day activities of the commercial bank. Hence, the bank holding company may have even greater demands than owners at stockholder banks to receive higher quality financial information from its subsidiaries for monitoring purposes.

Growth, profitability, and/or capital deficiency of a bank may also affect its financial reporting quality. We measure the bank's growth rate using change in total assets (GR), profitability using the return on average assets (ROAA), and the level of capital adequacy using total loans over total equity capital (LTC). The banks performance can affect accounting in two contrasting ways. First, a bank's accounting system may be incapable of handling excessive profitability and growth and thus hinder financial reporting quality. Second, an unprofitable bank or an undercapitalized bank may not be motivated or have the incentive to have accurate financial reporting information. Furthermore, a bank that is less capitalized or highly leveraged can be subjected to regulatory closure. Finally, (AI) is used in the model to indicate those banks that were audited. An independent audit should reduce the propensity to have a restatement and the magnitude of those restatements.

4.2.1. Restatement Baseline Models

The discussions presented above can be summarized into the following empirical models:

Model 1 – Likelihood of Restatements

RS_{iT+1} = PROBIT ($\alpha + \beta_1 LTA_{it} + \beta_2 OFF_{it} + \beta_3 NIITI_{it} + \beta_4 OBTA_{it} + \beta_5 MU_{it} + \beta_6 BHC_{it} + \beta_7 GR_{it} + \beta_8 LTC_{it} + \beta_9 ROAA_{it} + \beta_{10} AI_{it} + \varepsilon_{it}$)

Model 2 – Magnitude of Restatements

$RSABS_{iT+1}$ = GAMMA ($\alpha + \beta_1 LTA_{it} + \beta_2 OFF_{it} + \beta_3 NIITI_{it} + \beta_4 OBTA_{it} + \beta_5 MU_{it} + \beta_6 BHC_{it} + \beta_7 AI_{it} + \varepsilon_{it}$)

where,

RS	=	Restatement (Dummy Variable - 1 if the bank restated and 0 otherwise);
RSABS	=	Absolute value of Restatement;
LTA	=	Log of Total Assets;
OFF	=	Number of offices, branches, locations, and facilities;
NIITI	=	Non-interest Income divided by total Interest and Non-interest income;
OBTA	=	Off-balance sheet activities divided by Total Assets;

MU	=	Mutual or stockholder bank (Dummy Variable - 1 if a mutual bank and 0 a stock bank);
BHC	=	Parent is a Bank Holding Company (Dummy Variable - 1 if the bank's parent is a bank holding company and 0 otherwise);
GR	=	Change in Total Assets divided by beginning Total Assets;
LTC	=	Total Loans divided by Total Equity Capital;
ROAA	=	Net income (Loss) divided by Total Average Assets (assets at the end of the previous year plus assets at the end of the current year divided by 2); and
AI	=	Audit Indicator (Dummy Variable - 1 if the bank is independently audited and 0 otherwise).

4.2.2. Endogeneity/Selection Bias

The decision to have an independent audit may be endogenous at small private commercial banks. Hence, we may also have an endogeneity issue when analyzing the independent audit effect on the likelihood of a restatement **(Model 1)**. The decision to have an independent audit or not to have an audit is voluntarily made by the banks under study. Certain characteristics of a company may make management more likely to have an audit. These characteristics in turn may also affect the likelihood of a restatement. Thus, without controlling for the decision to have an audit or not to have an audit can cause our baseline model **(Model 1)** to be driven by these characteristics or the unobserved characteristics that affect the decision to have an audit. In order to overcome potential endogeneity issues, we use the Bi-variate Probit Regression to simultaneously model (Greene 2003; Poirier 1980) the decision to have an audit **(Model 6)** and the occurrence of a restatement **(Model 1)**. Bi-variate Probit Regression is used instead of Two Stage Least Square because **(Model 6)** and **(Model 1)** have dependent variables that are both binary. In Two Stage Least Square, the dependent variables in both models have to be continuous.

For analyzing the magnitude of restatements **(Model 2)**, we only examine banks with restatements. Hence, we may have selection bias issue where the dependent variable (magnitude of restatement) is only observed for a restricted non-random sample. In order to overcome the selection bias issue, we utilize Heckman's Selection Correction Model (Two-Step Estimation) (Heckman 1979). In the first stage, a Probit model **(Model 1)** is used to predict the probability of having a restatement and the Inverse Mills Ratio is calculated[14]. In the second stage, the Inverse Mills Ratio computed from Stage 1 is used in the restatement baseline model **(Model 2)** as an independent variable to control for selection bias and determine if we have a selection bias issue.

4.3. Discretionary Accruals

4.3.1. Accrual Estimation Model

We need to estimate the discretionary component of the allowance for loan losses account to analyze the effect that an independent audit has on conservative recognition of losses and the magnitude of discretionary accruals used in reporting. Both the nondiscretionary and discretionary component of the allowance for loan losses

[14] The Inverse Mills Ratio is the ratio of the probability density function (PDF) to the cumulative density function (CDF) for the predicted values from the restatement characteristic baseline model if there was a restatement, and the ratio of the PDF to (1 minus the CDF) if there was no restatement.

is unobservable and thus is require to be estimated. We use (Beaver and Engel 1996) Ordinary Least Squares Regression model (OLS) to estimate the nondiscretionary and discretionary component of the allowance for loan losses[15][16](**Model 3**).

Model 3 – Accrual Estimation

$DALL_{it} = OLS\ (\alpha_{it} + \beta_1 DCO_{it} + \beta_2 DTL_{it} + \beta_3 DNPL_{it} + \beta_4 D\Delta NPL_{it+1} + \varepsilon_{it})$,

Where,

i	=	Commercial bank identifier;
t	=	Year (2001 to 2010);
DALL	=	Allowance for loan losses;
DCO	=	Loan charge offs;
DTL	=	Total Loans;
DNPL	=	Nonperforming loans; and
ΔDNPL	=	Change in nonperforming loans as a percentage of the average of beginning and ending total loans.

In the (Beaver and Engel 1996) model, current charge offs (DCO) is used to provide some insight into the collectability of current and future loans. Nonperforming loans (DNPL) is used because it is an indication of assets that are in danger of default risk. Furthermore, one year ahead change in nonperforming loans (ΔDNPL) is used to proxy management's expectation of future defaults that is not reflected in the other explanatory variables. In addition to nonperforming loans, total loans (DTL) are used

[15] Each variable was deflated by "gross" book value of common equity (Beaver and Engel 1996).
[16] (Cook 1977) distance criterion was used to remove influential outliers.

because some default risk also exists in loans not designated as nonperforming. The predicted value from the model is defined as the nondiscretionary component and the error term or residual is defined as the discretionary component of the allowance for loan loss account.

4.3.2. Discretionary Accrual Baseline Models

Probit Regression (**Model 4**) is used to find the effect that an independent audit (AI) has on the conservative recognition of probable loan losses. Probit Regression is used because the dependent variable (conservative recognition of loan losses) is binary (Imai et al. 2007b). We use the sign of the residual from the (Beaver and Engel 1996) model (**Model 3**) as the dependent variable to proxy for conservative recognition of loan losses. A positive residual value means that the allowance for loan losses was over reported resulting in decreased income for the period (conservative). In contrast, a negative residual means that the allowance for loan losses was under reported resulting in increased income for the period (aggressive). In this study, we specifically define a company being more conservative in recognizing probable loan losses as having a positive residual. We use Gamma Regression (**Model 5**) to model the association of an independent audit with the magnitude of discretionary accruals. Gamma Regression is used because the dependent variable has a positive distribution (Imai et al. 2007a). The absolute value of the residual from the (Beaver and Engel 1996) model (Model 3) is used as the dependent variable to proxy for the magnitude of discretionary accruals used.

(Wahlen 1994) uses these control variables: beginning of period's total loan (SLTL), change in nonperforming loans (SΔNPL), beginning of period's nonperforming loans (SLNPL), and beginning of period's allowance for loan losses (SLALL)[17] to analyze the information each loan loss disclosure conveys to investors. Instead of focusing on investor's expectation of current and future loan losses, we model for management's expectation since it is unobservable to the public. We use similar control variables as (Wahlen 1994) in our study and include the audit indicator (AI) variable to analyze the effect that an independent audit has on the conservative reporting of the probable loan losses and the magnitude of discretionary accruals. Instead of breaking total loans into six loan categories like (Wahlen 1994), we use total loans for simplicity[18]. According to (Wahlen 1994), investors may use the beginning balance of total loans (SLTL) to form current period expectations of loan losses. Beginning non-performing loans (SLNPL) may be used to provide insight on current period nonperforming loans. Furthermore, the change in nonperforming loans (SΔNPL) can be used as a leading indicator of potential future loan losses. Lastly, the beginning balance of the allowance for loan losses (SLALL) is used as a measure of prior provisioning for loan losses. Investors may use past provisioning to form expectations on current and future provisions and chargeoffs.

[17] (Wahlen 1994) scaled the control variables using the beginning market value of equity. However, our study uses beginning book value of equity (except the audit indicator) because we are analyzing private banks.

[18] The results of using total loans instead of total loans broken into six loan categories are nearly the same.

Model 4 – Discretionary Accrual Conservatism

$rALLDirect_{it}$ = PROBIT ($\alpha + \beta_1 SLTL_{it} + \beta_2 S\Delta NPL_{it} + \beta_3\ SLNPL_{it-1} + \beta_4 SLALL_{\,it-1} + \beta_5 AI_{it} + \varepsilon$),

Model 5 – Discretionary Accrual Magnitude

$rALLABS_{it}$ = GAMMA ($\alpha + \beta_1 SLTL_{it} + \beta_2 S\Delta NPL_{it} + \beta_3\ SLNPL_{it-1} + \beta_4 SLALL_{\,it-1} + \beta_5 AI_{it} + \varepsilon$),

Where,

rALLDirect	=	Discretionary component of allowance for loan losses (Residual from (Beaver and Engel 1996) Model);
rALLABS	=	Magnitude of discretionary component of allowance for loan losses (Absolute value of residual from (Beaver and Engel 1996) Model;
SLTL	=	Beginning balance of total loans;
SΔNPL	=	Change in Nonperforming Loans;
SLNPL	=	Beginning balance of nonperforming loans;
SLALL	=	Beginning balance of allowance for loan losses; and
AI	=	Audit Indicator (Dummy Variable - 1 if the bank is independently audited and 0 otherwise).

4.3.3. Endogeneity

Like the restatement analysis above, we may have an endogeneity issue when analyzing the independent audit effect on the conservative recognition of losses (**Model 4**) and the magnitude of discretionary accruals (**Model 5**) used in financial reporting. The decision to have an independent audit or not to have an audit is voluntarily made by the banks in this study. Certain characteristics of a company may make management more likely to have an audit. These characteristics may also affect the decision to be conservative in recognizing probable loan losses and the magnitude of discretionary accruals used. Thus, without controlling for the decision to have an audit or not to have an audit can cause our baseline models (**Model 4** and **Model 5**) to be driven by these characteristics or the unobserved characteristics that affect the decision to have an audit. We use Bi-variate Probit Regression (Greene 2003; Poirier 1980) to simultaneously model the decision to have an audit **(Model 6)** and the conservative recognition of losses **(Model 4)**. Bi-variate Probit Regression is used because **Model 4** and **Model 5** both have binary dependent variables.

For the analysis of the magnitude of discretionary accruals, we use Simultaneous Probit - Gamma Regression **(Table 24)** to model the decision to have an audit **(Model 6)** and the magnitude of discretionary accruals **(Model 5)**. Simultaneous Probit - Gamma Regression is a derivative of the Residual Inclusion test proposed by (Wooldrige 1997). (Staub 2009) discusses the use of a Residual Inclusion test for testing endogeneity by simultaneously modeling two equations with one having a binary dependent variable and the other having a count dependent variable. (Staub 2009) finds the Residual Inclusion test performed as well as more complicated endogeneity test. We adopt this methodology since Gamma regression is similar to Poisson Regression. Poisson Regression is often used to model count data which is also positively distributed. In the Residual Inclusion test, we first model the decision to have an audit and generate the residuals. We then model the magnitude of discretionary accruals and include the residuals from modeling the decision to have an audit as an additional independent variable. If the residuals in the second model are significant then we may have an endogeneity issue. To the best of our knowledge, this is the first study to use it in the accounting and auditing literature to test for endogeneity.

Chapter 5

5. RESULTS & FINDINGS

5.1.Machine Learning Algorithms

H1: *The decision to have an independent audit is systematic and endogenous.*

First, using Weka's Experimenter **(Table 8)**, we compared the accuracy of six learning algorithms: Naïve Bayes (John and Langley 1995), Simple Logistic (Sumner et al. 2005), Support Vector Machine (SMO) (Hastie and Tibshirani 1998; Keerthi et al. 2001; Platt 1999), IBK (Aha et al. 1991), JRIP (Cohen et al. 2002), and Random Forest (Breiman 2001). Comparatively, the Simple Logistic (72.23 %), SMO (71.81 %), JRIP (71.92 %), and Random Forest (72.45 %) learning algorithm performed the best out of the six learning algorithms and have very similar accuracy rates in predicting whether a bank is audited or not audited **(Table 8)**[19]. We utilize a 2 sided T-Test to test whether these accuracy means are significantly different than 56%[20]. The T-test for each of the learning algorithms was highly significant and thus indicates that these four algorithms were able to predict with greater accuracy than 56%. The ability of these four algorithms to predict with accuracy above 70% indicates that the decision to have an audit is systematic and endogenous given the set of bank characteristics variables used. Thus, these four algorithms' performances are amenable to further analysis.

---

[19] The results are from a 10 fold cross validation of the entire dataset (pooled). We also ran a 10 fold cross validation using only a specific year and the results were materially similar.

[20] 56% of banks decided to have an audit and 44% did not have an audit.

Table 8

Weka Experimenter Results

Machine Learning Algorithms	**Accuracy Rate**
NaïveBayes[21]	46.52 %
SimpleLogistic[22]	72.23 %
SMO[23]	71.81 %
IBK[24]	66.57 %
JRIP[25]	71.92 %
RandomForest[26]	72.45 %

We look further into the Simple Logistic **(Table 9)**, SMO **(Table 10)**, JRIP **(Table 11)**, and Random Forest **(Table 12)** learning algorithms to examine their prediction performance in greater detail using Weka's Explorer. In Weka's Explorer program, the overall accuracy rates were very similar to Weka's Experimenter results (Simple Logistic (72.2793 %), SMO algorithms (71.8074%), JRIP (71.6854%), and Random Forest (72.61%)). The slight difference in accuracy rates is attributed to Weka's Explorer running a single 10 fold cross-validation by default, whereas Weka's Experimenter runs 10 runs of 10-fold cross validation by default. We explore further with other performance measurements using True Positive, False Positive, Precision, Recall, F-Measure, and ROC rates.

---

[21] Parameters: bayes.NaiveBayes '' 5995231201785697655
[22] Parameters: functions.SimpleLogistic '-I 0 -M 500 -H 50 -W 0.0' 7397710626304705059
[23] Parameters: functions.SMO '-C 1.0 -L 0.0010 -P 1.0E-12 -N 0 -V -1 -W 1 -K \"functions.supportVector.PolyKernel -C 250007 -E 1.0\"' -6585883636378691736
[24] Parameters: lazy.IBk '-K 1 -W 0 -A \"weka.core.neighboursearch.LinearNNSearch -A \\\"weka.core.EuclideanDistance -R first-last\\\"\"' -3080186098777067172
[25] Parameters: rules.JRip '-F 3 -N 2.0 -O 2 -S 1' -6589312996832147161
[26] Parameters: trees.RandomForest '-I 10 -K 0 -S 1' 4216839470751428698

The True Positive rate and False Positive rate shows the individual class prediction accuracy and support the overall accuracy of the algorithms. A high True Positive rate indicates that the algorithm was good at predicting the right class (Audited vs. Unaudited). The True Positive rates (TP Rate) for predicting the audited class was as follows: Simple Logistic (0.79), SMO (0.792), JRIP (0.77), and RandomForest (0.737) and the True Negative rates for predicting the unaudited class were as follows: Simple Logistic (0.636), SMO (0.622), JRIP (0.648), and RandomForest (0.713). The four learning algorithms also have low False Positive rates (FP Rate) when predicting either class (Audited/Unaudited) which further supports the high accuracy rates. A low False Positive rate indicates that the algorithm seldom predicted the wrong class. The False Negative rates for predicting the audited class was as follows: Simple Logistic (0.364), SMO (0.378), JRIP (0.353), and RandomForest (0.287) and the False Positive rates for predicting the unaudited class were as follows: Simple Logistic (0.21), SMO (0.208), JRIP (0.23), and RandomForest (0.263).

We also get confirming performance measurements from the Precision, Recall, F-Measure, and ROC rates. The Precision rates for predicting the audited class was as follows: Simple Logistic (0.738), SMO (0.731), JRIP (0.74), and RandomForest (0.769), and the Precision rates for predicting the unaudited class were as follows: Simple Logistic (0.7), SMO (0.697), JRIP (0.684), and RandomForest (0.675). The high Precision rates indicate that the four learning algorithms have a low number of false positive and indicate the accuracy of the algorithms in predicting the correct class. The Recall rates for predicting the audited class was as follows: Simple Logistic (0.79), SMO (0.792), JRIP (0.77), and RandomForest (0.737), and the Recall rates for predicting the unaudited class were as follows: Simple Logistic (0.636), SMO (0.622), JRIP (0.684), and RandomForest (0.713). A higher Recall rate indicates that the learning algorithms had few misclassifications. Moreover, a high F-Measure confirms the high Precision and Recall rate. The F-Measure rates for predicting the audited class was as follows: Simple Logistic (0.763), SMO (0.761), JRIP (0.755), RandomForest (0.752) and the F-Measure rates for predicting the unaudited class were as follows: Simple Logistic (0.666), SMO (0.657), JRIP (0.666), and RandomForest (0.694). Lastly, the high ROC rates show that the learning algorithms performed better than making a random guess. A ROC rate around .50 indicates that the model is no better than a random guess (Tan et al. 2005). The ROC Area rates for predicting the audited class was as follows: Simple Logistic (0.793), SMO (0.707), JRIP (0.724), and RandomForest (0.794) and the ROC rates for predicting the unaudited class were as follows: Simple Logistic (0.793), SMO (0.707), JRIP (0.724), and RandomForest (0.794). Hence, the four algorithms prediction ability is superior to a random guess and further supports the decision to have an audit is systematic and endogenous given the set of bank characteristics variables used.

Table 9

Simple Logistic Algorithm
Weka Explorer Results

Percent Correctly Classified Instances: 72.2793%			**Percent Incorrectly Classified Instances:** 27.7207%			
Predicted Class / **Class**	**0**	**1**	**Precision**	**Recall**	**F-Measure**	**ROC Area**
0	0.636	0.364	0.7	0.636	0.666	0.793
1	0.21	0.79	0.738	0.79	0.763	0.793

Table 10

SMO Algorithm
Weka Explorer Results

Percent Correctly Classified Instances: 71.8074%			**Percent Incorrectly Classified Instances:** 28.1926%			
Predicted Class / **Class**	**0**	**1**	**Precision**	**Recall**	**F-Measure**	**ROC Area**
0	0.622	0.378	0.697	0.622	0.657	0.707
1	0.208	0.792	0.731	0.792	0.761	0.707

Table 11

JRIP Algorithm
Weka Explorer Results

Percent Correctly Classified Instances: 71.6854%			**Percent Incorrectly Classified Instances:** 28.3146%			
Predicted Class / **Class**	**0**	**1**	**Precision**	**Recall**	**F-Measure**	**ROC Area**
0	0.648	0.352	0.684	0.684	0.666	0.724
1	0.23	0.77	0.74	0.77	0.755	0.724

Table 12

RandomForest Algorithm
Weka Explorer Results

Percent Correctly Classified Instances: 72.61%			**Percent Incorrectly Classified Instances:** 27.39%			
Predicted Class / **Class**	**0**	**1**	**Precision**	**Recall**	**F-Measure**	**ROC Area**
0	0.713	0.287	0.675	0.713	0.694	0.794
1	0.263	0.737	0.769	0.737	0.752	0.794

Unlike tradition statistical outputs, the interpretation of machine learning algorithm outputs is often difficult to decipher or the outputs are nonexistent. However, the output from the Simple Logistic machine learning algorithm has similar output characteristics as the traditional generalized linear model (logistic regression) and can be used to approximately capture the “first order” effects of the independent variables used on the decision to have a voluntary independent audit. **Table 13** shows the output from the Simple Logistic learning algorithm. For interpretation purposes, we focus on Class 1. Class 1 is the audited class. As expected and in line with the literature, larger (LTA – Positive Coefficient (0.54)), growing (GR – Positive Coefficient (0.05)), and less profitable banks (ROAA – Negative Coefficient (-0.2)) are more likely to have an independent audit absent regulation. Furthermore, banks that have more complex operations (NIITI - Positive Coefficient (0.23), OBTA – Positive Coefficient (2.57)) also has a higher likelihood of procuring an independent audit. We also find that hierarchical or ownership structure may affect the decision to have an audit. More specifically, banks that are part of bank holding company (BHC - Negative Coefficient (-0.25)) are less likely to have an audit and banks owned by depositors (MU - Positive Coefficient (2.36)) are more likely to have an audit. This finding is in contradiction to expectation. We expected that commercial banks that are part of a bank holding company and stockholder banks would demand an independent audit over non-bank holding companies and mutual banks, respectively. Collectively, these results suggest that profitable and growing banks with complex operations have a higher likelihood of procuring an independent audit. Furthermore, hierarchal and ownership structure may influence the decision to have or not have an audit.

Table 13

Simple Logistic Algorithm
Weka Explorer Results

Parameter:	**Class 0 (Unaudited):**	**Class 1 (Audited):**
Intercept	-0.42	0.42
LTA	-0.54	0.54
NIITA	-0.23	0.23
OBTA	-2.57	2.57
GR	-0.05	0.05
ROAA	0.2	-0.2
MU	-2.36	2.36
BHC	0.25	-0.25

Note: All variables significant (95% level)

5.1.1. Robustness Test

As a robustness test, we run traditional Probit Regression in SAS to see if the Simple Logistic machine learning algorithm output results are comparable. Regression type models are commonly used in the accounting and auditing literature. Our Probit Regression model is presented below (**Model 6**):

Model 6 – Audit Decision Model

$AI_{it} = PROBIT(\alpha_{tr} + \beta_1 LTA_{it} + \beta_2 NIITI_{it} + \beta_3 OBTA_{it} + \beta_4 GR_{it} + \beta_5 ROAA_{it} + \beta_6 MU_{it} + \beta_7 BHC_{it} + \varepsilon_{it})$,

Where,

i	=	Commercial bank identifier;
t	=	Year (2001 to 2010);
AI	=	Audit Indicator (Dummy Variable - 1 if the bank is independently audited and 0 otherwise);
LTA	=	Log of total assets;
NIITI	=	Non-interest income divided by total income;
OBTA	=	Off-balance sheet activities divided by total assets;
GR	=	Change in total assets divided by beginning total assets;
ROAA	=	Net income divided by average assets;
MU	=	Mutual or stockholder bank (Dummy Variable - 1 if a mutual bank and 0 a stock bank); and
BHC	=	Parent is a Bank Holding Company (Dummy Variable - 1 if the bank's parent is a bank holding company and 0 otherwise).

According to **Table 14**, the SAS Probit Regression coefficients are very comparable to Weka's Simple Logistic coefficients **(Table 13)**. We also compare the confusion matrix from the Weka's Simple Logistic Algorithm and from SAS's Probit Regression model to compare the classification accuracy of both models **(**

Table 15**)**. Both models yield similar True Positive and False Positive rates. For example, True Positive rates for the audited class in Weka is 79% and in SAS is 73.84% and for the unaudited class in Weka is 63.6% and in SAS is 69.5%. The difference in the accuracy for both Logistic Regression models is mainly attributed to Weka using the LogitBoost algorithm and SAS using Fisher's scoring algorithm.

Table 14

Audit Decision – Logistic Regression Model

Parameter	Estimate	Standard Error	Wald Chi Square	Pr > ChiSq
Intercept	0.4765	0.0259	339.3838	<.0001
LTA	0.6405	0.00895	5117.153	<.0001
NIITI	0.2408	0.0103	544.9048	<.0001
OBTA	2.7192	0.4795	32.1654	<.0001
GR	0.0549	0.0127	18.6792	<.0001
ROAA	-0.1764	0.0132	178.328	<.0001
MU	2.2497	0.1208	347.0683	<.0001
BHC	-0.2883	0.0231	155.56	<.0001
R-Square	0.2371			
Max-rescaled R-Square	0.3179			

Table 15

Confusion Matrix

	Weka Confusion Matrix		SAS Confusion Matrix	
	0 - Unaudited	1 - Audited	0 - Unaudited	1 - Audited
0 - Unaudited	8,610 63.6%	4,938 21%	8,625 69.95%	3,705 30.05%
1 - Audited	3,697 36.4%	13,905 79%	4,923 26.16%	13,897 73.84%

5.2.Restatement Models

H2: *An independent audit decreases the likelihood of having restatements.*

H3: *An independent audit decreases the magnitudes of restatements.*

5.2.1. Restatement Baseline Models

According to **Table 17B**, we have endogeneity issue at the 90% confidence level (Rho (P-value 0.066)). Hence, we will focus our analysis on the results from the Bi-variate Probit Regression **(Table 17B)** model instead of the Probit Regression Model **(Table 16)**. The results from modeling the decision to have an audit in the Bi-variate Probit Regression model are presented in **Table 17A**. We expected that audited banks should have a lower propensity to have a restatement when compared with unaudited banks. In **Table 17B**, the coefficient estimate for the audit indicator variable (AI) was positive (0.343049) which suggests statistically that banks who have an audit were more likely to have restatements. The finding is counter intuitive as we expected that audited banks would have a lesser likelihood of having a restatement. For analyzing the magnitude of restatements, we focus our analysis on **Table 19B** (Heckman's Selection Correction Model) instead of **Table 18** (Gamma Regression Model) since the Inverse Mills Ratio variable (IMR) is significant (P-value < 0.0001) and hence we have a selection bias problem. The likelihood of a restatement **Table 19A** is modeled in the first stage of the Heckman Correction Model. In line with expectations, the coefficient estimate for the audit indicator variable **(Table 19B)** shows that audited banks were more likely to have restatements with lower magnitudes (AI – Negative Coefficient (-0.2895)). Although the finding that audited banks have restatements with lower magnitudes is positive for auditors, the fact that a restatement was necessary in the first place still indicates that a quality audit was not performed.

Table 16

Likelihood of Restatement – Probit Regression Model

Model 1 – Likelihood of Restatements

Parameter	Estimate	Standard Error	WaldChi Square	Pr > ChiSq
Intercept	-1.5229	0.0369	1704.157	<.0001
LTA	0.0434	0.0152	8.1481	0.0043
OFF	0.0132	0.00414	10.1132	0.0015
NIITI	0.00658	0.012	0.3002	0.5838
OBTA	0.0309	0.00918	11.3589	0.0008
MU	0.1699	0.0579	8.6039	0.0034
BHC	-0.119	0.0323	13.5278	0.0002
GR	0.00749	0.00843	0.7883	0.3746
LTC	-0.0485	0.016	9.2126	0.0024
ROAA	-0.0621	0.0135	21.2011	<.0001
AI	0.0601	0.0259	5.3931	0.0202
R-Square	0.0046			
Max-rescaled R-Square	0.012			

Table 17

Likelihood of Restatement – Bivariate Probit Model (Endogeneity)

A. Model 6 – Audit Decision Model

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	2.43301	0.122911	19.79	<.0001
LTA	0.64038	0.008887	72.06	<.0001
NIITI	0.241683	0.00966	25.02	<.0001
OBTA	2.679241	0.498078	5.38	<.0001
GR	0.055148	0.012421	4.44	<.0001
ROAA	-0.1778	0.010619	-16.74	<.0001
MU	2.245994	0.121508	-18.48	<.0001
BHC	-0.28875	0.023095	12.5	<.0001

B. Model 1 – Likelihood of Restatements

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	-1.36664	0.065349	-20.91	<.0001
LTA	-0.01352	0.034466	-0.39	0.6948
OFF	0.012054	0.004182	2.88	0.0039
NIITI	-0.01366	0.016664	-0.82	0.4122
OBTA	0.028204	0.009538	2.96	0.0031
MU	0.067658	0.08464	-0.8	0.4241
BHC	-0.09185	0.03568	2.57	0.01
GR	0.005188	0.00864	0.6	0.5482
LTC	-0.05056	0.015761	-3.21	0.0013
ROAA	-0.05196	0.014061	-3.7	0.0002
AI	0.343049	0.155709	-2.2	0.0276
Rho	-0.17341	0.094318	-1.84	0.066

Table 18

Restatement Magnitude – Gamma Regression Model

Model 2 – Magnitude of Restatements

Parameter		Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		4.5197	0.0932	2350.04	<.0001
LTA		0.7166	0.0363	389.9	<.0001
OFF		-0.0421	0.009	21.81	<.0001
NIITI		0.1043	0.0358	8.48	0.0036
OBTA		0.0265	0.0178	2.21	0.137
MU	1	0.8873	0.1422	38.92	<.0001
BHC	1	0.3352	0.0853	15.43	<.0001
AI	1	-0.0097	0.0678	0.02	0.8859
Scale		0.5739	0.0152		

Criteria For Assessing Goodness Of Fit

Criterion	Value/DF
Deviance	2.1764

Table 19

Restatement Magnitude – Heckman Selection Correction Model (Selection Bias)

A. Model 1 – Likelihood of Restatements

Parameter	Estimate	Standard Error	WaldChi Square	Pr > ChiSq
Intercept	-1.5229	0.0369	1704.157	<.0001
LTA	0.0434	0.0152	8.1481	0.0043
OFF	0.0132	0.00414	10.1132	0.0015
NIITI	0.00658	0.012	0.3002	0.5838
OBTA	0.0309	0.00918	11.3589	0.0008
MU	0.1699	0.0579	8.6039	0.0034
BHC	-0.119	0.0323	13.5278	0.0002
GR	0.00749	0.00843	0.7883	0.3746
LTC	-0.0485	0.016	9.2126	0.0024
ROAA	-0.0621	0.0135	21.2011	<.0001
AI	0.0601	0.0259	5.3931	0.0202

B. Model 2 – Magnitude of Restatements

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	13.9335	1.3915	100.27	<.0001
LTA	0.6062	0.0395	235.6	<.0001
OFF	-0.0965	0.0121	63.35	<.0001
NIITI	0.1266	0.0359	12.46	0.0004
OBTA	-0.0657	0.0237	7.71	0.0055
MU	0.1217	0.1816	0.45	0.5029
BHC	0.8422	0.1117	56.81	<.0001
AI	-0.2895	0.0796	13.22	0.0003
IMR	-4.8301	0.7099	46.29	<.0001
Scale	0.5835	0.0154		

Criterion	Value/DF
Deviance	2.1363

For the control variables, we expected larger banks to have a lower likelihood of having a restatement. The results in **Table 17B** confirm this by showing that larger banks (LTA – Positive Coefficient (-0.01352)) have a lower likelihood of having a restatement. Although the results are insignificant for LTA in **Table 17B**, we are analyzing the likelihood of a restatement (**Table 17B**) simultaneously with the decision to have an audit (**Table 17A**) and must also consider the significance of LTA in **Table 17A**. In **Table 17A**, LTA is significant and hence is significant in the model. It can be argued that larger banks are involved in more complex operations and thus may have more complicated accounting. As a result, larger banks may have more restatements than smaller banks. However, to compensate, larger banks typically have better internal controls and managers with greater financial expertise than smaller banks.

We expected banks with a greater number of branches/offices and that are involved in nontraditional banking products and services to have more complex accounting and thus have a higher likelihood of a restatement. In **Table 17B**, the Bivariate Probit Regression's coefficient estimate for the variable OFF was 0.012054 which indicates that banks with more branches/offices are associated with more restatement. For banks with greater complexity in operations, we find that banks with greater amounts of non-interest income (NIITI – Negative Coefficient (-0.01366)) were less likely to have restatements while banks with greater off balance sheet activity (OBTA – Positive Coefficient (0.028204)) were more likely to have a restatement. NIITI was not significant in **Table 17B** but since we are analyzing the likelihood of a restatement (**Table 17B**) simultaneously with the decision to have an audit (**Table 17A**), we must also consider the significance of NIITI in **Table 17A**. In **Table 17A**, NIITI is significant. The result for NIITI is contradictory to what we expected. We expected a bank with more complex non-interest products and services to have more restatements. However, perhaps banks with greater NIITI are typically bigger banks and they may have better internal controls to reduce material misstatements.

For ownership type, the results indicate that a mutual bank (MU – Positive Coefficient (0.067658)) was more likely to have a restatement than a stockholder bank. The finding is consistent with our expectation. Unlike depositors/owners of mutual banks, the owners/managers at stockholder banks may be involved in management activities and hence have a demand for higher quality financial information as a monitoring mechanism. Depositors/owners of mutual banks may not care about the performance of the mutual bank since their funds are federally insured. Likewise for hierarchical type, bank holding companies may have similar financial reporting quality demands as owners/managers of stockholder banks. However, unlike owners/managers at stockholder banks, the bank holding company may be absent from the day to day operations of the commercial bank subsidiary. Thus, a bank holding company may also demand higher quality financial information as a governance mechanism to encourage management at their subsidiaries to use more stewardship in management and accounting. The results are consistent with this expectation and show that commercial banks that are part of a bank holding company (BHC – Negative Coefficient (-0.09185)) were less likely to have a restatement.

Our findings in **Table 17B** indicate that growing banks (GR – Positive Coefficient (0.005188)) have a higher likelihood of a restatement. The results for GR in **Table 17B** are not significant but we must consider the significant results for GR from **Table 17A**. Our finding that higher growth may increase the likelihood of restatements is consistent with what we expected and a potential explanation could be that the accounting system at growing banks may break down due to their inability to keep up with the growth of the bank and thus increase the likelihood of a restatement. In addition, we expected that management at banks with profitability and/or capital adequacy issues may not be motivated to have more accurate and reliable financial information. In fact, management may have the inclination to prop financial performance to meet growth or profitability expectations and regulatory capital requirements. The results in **Table 17B** show that less profitable (ROAA – Negative Coefficient (-0.05196)) and undercapitalized banks (LTC – Negative Coefficient (-0.05056)) are more likely to have restatements. These finding perhaps suggest that auditors and regulators should pay more attention to the financial reporting of less profitable and undercapitalized banks during their examinations.

5.3.Discretionary Accrual Models

***H4:** An independent audit increases the conservative recognition of expected and probable losses.*

***H5:** An independent audit decreases the magnitudes of discretionary accruals in financial reporting.*

5.3.1. Accrual Estimation Model

Table 20 shows the results from the (Beaver and Engel 1996) accrual estimation model. The objective of the (Beaver and Engel 1996) model is to estimate the nondiscretionary and discretionary component of the allowance for loan losses for subsequent analysis since they are both unobservable. The residual generated by the accrual estimation model is defined as the estimated discretionary component of the allowance for loan losses balance. Moreover, the fitted value generated by the accrual estimation model is defined as the estimated nondiscretionary component of the allowance for loan losses balance. The overall model's explanatory power (Adjusted R-Square) is 43% and the control variables used are all significant in the model except for change in nonperforming loans (ΔDNPL). In line with expectations, the significant positive parameter coefficients for loan charge offs (DCO Coefficient = 0.29423), nonperforming loans (DNPL Coefficient = 0.24029), and total loans (DTL Coefficient = 0.29423) indicate as the quality of the loan portfolio deteriorates and more loans are made then the higher the allowance for loan losses balance will be. This is consistent with expectations because the allowance for loan loss account may need to increase in order to compensate for an increase in probable credit losses. Although change in nonperforming loans was not significant, the coefficient was positive (ΔDNPL Coefficient = 0.00518).

Table 20

Beaver Model – OLS Regression Model (Accrual Estimation)

Model 3 – Accrual Estimation

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1.40E-14	0.00428	0	1
DCO	0.29423	0.00486	60.59	<.0001
DTL	0.38099	0.00444	85.87	<.0001
DNPL	0.24029	0.005	48.04	<.0001
ΔDNPL	0.00518	0.0044	1.18	0.2385
R-Square	0.4306			
Adj R-Sq	0.4306			

5.3.2. Baseline Models

The base line and endogeneity model results for analyzing the conservatism in reporting of probable credit losses are presented in **(Table 21)** and **(Table 22)**, respectively. Based on the significance of Rho **(Table 22B)** in the Bivariate Probit Regression model, we do have an endogeneity problem and thus focus our analysis on this model. The decision to have an audit in the Bivariate Probit Regression model is presented in **Table 22A**. The results indicate that audited banks (AI – Negative

Coefficient (-0.12302)) are less conservative in recognizing probable loan losses. These results show that audited banks have a tendency to over-estimated income and over-value their assets. According to (Basu 1997), conservatism is the recognition of losses on a more timely basis than gains or income. For analyzing the magnitude of discretion accruals, we focus our analysis in **Table 24B** instead of **Table 23** since the Residual Inclusion variable (RI) is significant and hence we have an endogeneity issue. The results from modeling the decision to have an audit in the Simultaneous Probit Gamma Regression model are presented in **Table 24A**. For the magnitude of discretion accruals, the results on **(Table 24B)** indicate that audited banks (AI – Positive Coefficient (0.196)) had higher magnitudes of discretionary accruals than unaudited banks. This shows that audited banks exposed themselves to potential estimation errors in using higher magnitudes of discretionary accruals in reporting. Collectively, these two findings show that audited banks have a tendency to underestimate probable loan losses and use higher magnitudes of discretionary accruals in financial reporting than unaudited banks. These two findings suggest that an independent audit does not promote financial reporting quality based on these specific measurements of financial reporting quality.

Table 21

Discretionary Accruals Conservatism – Probit Regression Model

Model 4 – Discretionary Accrual Conservatism

Parameter	Estimate	Standard Error	WaldChi Square	Pr > ChiSq
Intercept	-0.1178	0.0113	108.3031	<.0001
SLTL	-0.2235	0.0162	189.7086	<.0001
SΔNPL	-0.1866	0.0191	95.1639	<.0001
SLNPL	-0.2682	0.0114	550.0647	<.0001
SLALL	1.0131	0.0184	3027.907	<.0001
AI	-0.0677	0.0151	20.0369	<.0001
R-Square	0.2087			
Max-rescaled R-Square	0.2798			

Table 22

Discretionary Accruals Conservatism – Bivariate Probit Model (Endogeneity)

A. Model 6 – Audit Decision Model

Parameter Estimates	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	2.452713	0.123113	19.92	<.0001
LTA	0.640458	0.008898	71.98	<.0001
NIITI	0.258345	0.010056	25.69	<.0001
OBTA	2.657986	0.494232	5.38	<.0001
GR	0.051791	0.012353	4.19	<.0001
ROAA	-0.17889	0.010661	-16.78	<.0001
MU	2.266356	0.121758	-18.61	<.0001
BHC	-0.29037	0.023086	12.58	<.0001

B. Model 4 – Discretionary Accrual Conservatism

Parameter Estimates	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	-0.14103	0.018446	-7.65	<.0001
SLTL	-0.86313	0.021107	-40.89	<.0001
SΔNPL	-0.23996	0.019141	-12.54	<.0001
SLNPL	-0.55313	0.012712	-43.51	<.0001
SLALL	2.538942	0.029079	87.31	<.0001
AI	-0.12302	0.037339	3.29	0.001
Rho	0.076448	0.025433	3.01	0.0026

For the control variables, the results in **Table 22B** indicate that as nonperforming loans (SΔNPL and SLNPL) increase or become greater (quality of loan portfolio deteriorates) then the less conservative banks are in the reporting of potential loan losses. This is not in line with expectations. In addition, the magnitude of that un-conservatism is greater **(Table 24B)**. We expected that as nonperforming loans increase then the bank's becomes more conservative in reporting probable loan losses because potential credit losses are manifesting. In essence, the banks are being less conservative in reporting probable credit losses. Furthermore, the size of the loan portfolio (SLTL) did not seem to play a role in influencing conservatism. In line with the immediate previous explanation for nonperforming loans, we expected the larger a bank's loan portfolio gets the more concern the bank will be with reporting more probable loan losses.

Table 23

Discretionary Accruals Magnitude – Gamma Regression Model

Model 5 – Discretionary Accrual Magnitude

Parameter	Estimate	Standard Error	Wald Chi Square	Pr > ChiSq
Intercept	-0.6815	0.0076	7998.08	<.0001
SLTL	-0.0626	0.0093	45.04	<.0001
SΔNPL	0.1596	0.0115	193.44	<.0001
SLNPL	0.1208	0.0069	302.44	<.0001
SLALL	0.2666	0.0071	1390.93	<.0001
AI	0.0019	0.0102	0.04	0.8502
Scale	1.28	0.0092		

Criterion	Value/DF
Deviance	0.8781

Table 24

Discretionary Accruals Magnitude – Simultaneous Probit - Gamma Regression Model

A. Model 6 – Audit Decision Model

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	2.43301	0.122911	19.79	<.0001
LTA	0.64038	0.008887	72.06	<.0001
NIITI	0.241683	0.00966	25.02	<.0001
OBTA	2.679241	0.498078	5.38	<.0001
GR	0.055148	0.012421	4.44	<.0001
ROAA	-0.1778	0.010619	-16.74	<.0001
MU	2.245994	0.121508	-18.48	<.0001
BHC	-0.28875	0.023095	12.5	<.0001

B. Model 5 – Discretionary Accrual Magnitude

Parameter	Estimate	Standard Error	Wald Chi Square	Pr > ChiSq
Intercept	-0.7892	0.0148	2838.59	<.0001
SLTL	-0.0814	0.0096	72.03	<.0001
SΔNPL	0.1537	0.0115	179.85	<.0001
SLNPL	0.1172	0.007	284.03	<.0001
SLALL	0.2819	0.0074	1451.52	<.0001
AI	0.196	0.0252	60.42	<.0001
RI	-0.1004	0.0119	70.89	<.0001
Scale	1.2823	0.0092		

Criterion	Value/DF
Deviance	0.8764

5.4.Robustness Test

The findings discussed above result from examining banks that persistently had an audit or did not persistently had an audit for all 10 years under study. We used only banks that were persistent in their audit decision because the Call Reports are not directly audited by the independent auditors. However, the data used to generate the Call Reports were audited. Hence, we only look to see if an independent audit had any effect on financial reporting quality. Furthermore, the results in the study are astonishing as they contradict anecdotal evidence and intuition. Therefore, a robustness test of our findings is beneficial to support our findings and to provide additional insights. As a robustness test, we use the complete dataset of commercial banks to test whether independently audited banks were less likely to have restatements, have lower magnitudes of restatements, more conservative in recognizing probable loan losses, and more conservative in the use of discretionary accruals. The additional analysis of banks that had inconsistent audits can provide additional insight. In the complete dataset, we have banks that persistently had an audit or did not persistently had an audit, banks that had an audit in the prior period but did not have an audit in the current period, and banks that did not have an audit in the prior period but had an audit in the current period. Four dummy audit indicator variables are constructed to represent these situations.

CAI1	=	Audit Indicator (Dummy Variable – 1 if the bank is independently audited in the prior period and independently audited in the current period and 0 otherwise);
CAI2	=	Audit Indicator (Dummy Variable – 1 if the bank is not independently audited in the prior period and independently audited in the current period and 0 otherwise);
CAI3	=	Audit Indicator (Dummy Variable – 1 if the bank is independently audited in the prior period and not independently audited in the current period and 0 otherwise); and
CAI4	=	Audit Indicator (Dummy Variable – 1 if the bank is not independently audited in the prior period and not independently audited in the current period and 0 otherwise).

Unlike the main analysis, in the robustness test we must use a Multinomial Probit – Probit Regression model instead of a Bi-variate Probit Model. For analyzing the decision to have an audit in the robustness test, we have an audit indicator variable with four categories of audit procurement. As a result, in the first model we use a Multinomial Probit Regression to model the decision to have an audit instead of Probit

Regression. Multinomial Probit Regression is used when the dependent variable is categorical and to predict the probable outcome of each category based on a given set of independent variables (Greene 2003). In the Multinomial Probit – Probit Regression model, we chose the decision not to have an audit in the prior period and in the current period (CAI4) as the reference category and thus omitted from the analysis. The residuals (RI1, RI2, RI3) from Multinomial Probit Regression are included in the second model as independent variables. The inclusion of the residuals in the second model is used as a Residual Inclusion test for endogeneity testing (Wooldrige 1997).

We first look at the likelihood of a restatement and its magnitude. Based on the significance of the residual inclusion RI1 in **Table 26D** and IMR **i**n **Table 28B**, we do have an endogeneity and selection bias issue, respectively. Although RI2 and RI3 in **Table 26D** are also significant, we are primarily focus on the analysis of the consistent decision to have an audit. Hence, we focus on the results from the Multinomial Probit – Probit Regression model **(Table 26)** and the Heckman Selection Correction Model **(Table 28)** instead of the results from the Probit Regression (**Table 25**) and Gamma Regression (**Table 27**) models. **Table 26 A, B, C** presents the results from modeling the decision to have an audit for the four different categories. The residuals from these models are used as independent variables in the modeling of the likelihood of a restatement **(Table 26D)**. Furthermore, **Table 28A** presents the results from modeling the likelihood of a restatement and generating the inverse mills ratio (IMR) for the second stage **(Table 28B)** of the Heckman Selection Correction Model.

In **Table 26D**, the results confirm the finding that banks that had an independent audit consistently (CAI1) had a higher likelihood of having a restatement. Furthermore, we find that banks had a higher probability of a restatement if they did not have an audit in the prior period but had an audit in the current period. This finding is consistent with expectation and with the literature. (Lazer et al. 2004) find companies who switch to a new auditor have a higher likelihood of having a restatement. The new auditor may impose a restatement in the year of the audit in order to correct material misstatements and to limit their potential litigation risk going forward. The results from **Table 28B** also confirm the findings in the main study where consistently independently audited banks (CAI1) are found to have restatements with lower magnitudes. Furthermore, the results show significance for those banks who switch from not having an audit in the prior period to having an audit (CAI2) in the current period have lower restatement magnitudes as well. The finding for CAI2 is counterintuitive as we would expect that the magnitude of restatements would be larger for those banks that did not have an audit in the prior year but had an audit in the current year.

Table 25

Likelihood of Restatements – Robustness Test

Model 1 – Likelihood of Restatements

Parameter		Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept		-1.4888	0.0324	2117.838	<.0001
LTA		0.0281	0.0133	4.4976	0.0339
OFF		0.0127	0.00385	10.8931	0.001
NIITI		0.0224	0.0104	4.6746	0.0306
OBTA		0.0232	0.00808	8.2017	0.0042
MU	1	0.1722	0.0543	10.065	0.0015
BHC	1	-0.1301	0.0284	20.9912	<.0001
GR		0.00456	0.00808	0.3185	0.5725
LTC		-0.0468	0.0136	11.8628	0.0006
ROAA		-0.0616	0.012	26.4229	<.0001
CAI1	1	0.048	0.0229	4.3933	0.0361
CAI2	2	0.2363	0.0613	14.8769	0.0001
CAI3	3	0.0663	0.0671	0.9757	0.3233

R-Square	0.0043
Max-rescaled R-Square	0.011

Table 26

Likelihood of Restatements – Robustness Test – Multinomial Probit – Probit Regression Model (Endogeneity)

A. Model 6 – Audit Decision Model (Consistently Audited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	0.2989	0.022	184.1322	<.0001
LTA	0.562	0.0077	5332.011	<.0001
NIITI	0.2043	0.00856	570.0939	<.0001
OBTA	2.4823	0.4261	33.936	<.0001
GR	0.00361	0.00748	0.2329	0.6294
ROAA	-0.1567	0.0108	211.1812	<.0001
MU	1.7186	0.0758	514.3878	<.0001
BHC	-0.2337	0.0202	133.5625	<.0001

B. Model 6 – Audit Decision Model (Unaudited to Audited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	-2.2577	0.0818	761.3198	<.0001
LTA	-0.0806	0.0149	29.2916	<.0001
NIITI	-0.0159	0.0162	0.9565	0.3281
OBTA	-6.7588	2.4569	7.5677	0.0059
GR	0.0345	0.00879	15.3924	<.0001
ROAA	-0.0489	0.0186	6.9308	0.0085
MU	-0.3916	0.1263	9.608	0.0019
BHC	0.0777	0.0422	3.3903	0.0656

C. Model 6 – Audit Decision Model (Audited to Unaudited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	-2.1246	0.0716	879.4655	<.0001
LTA	-0.1267	0.0147	74.2202	<.0001
NIITI	-0.0147	0.016	0.8501	0.3565
OBTA	-3.0475	2.1007	2.1046	0.1469
GR	0.0192	0.00822	5.4495	0.0196
ROAA	-0.0222	0.0187	1.4025	0.2363
MU	-0.3654	0.1223	8.9258	0.0028
BHC	0.0339	0.0409	0.6883	0.4067

D. Model 1 – Likelihood of Restatements

Parameter		Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept		-2.4394	0.2363	106.5479	<.0001
LTA		0.00588	0.0447	0.0173	0.8953
OFF		0.0105	0.00388	7.3862	0.0066
NIITI		0.00418	0.0181	0.0532	0.8176
OBTA		0.0257	0.00874	8.6279	0.0033
MU	1	0.1411	0.1106	1.6275	0.2021
BHC	1	-0.1236	0.0344	12.9179	0.0003
GR		-0.0297	0.0118	6.3692	0.0116
LTC		-0.0423	0.0137	9.4933	0.0021
ROAA		-0.0247	0.0141	3.057	0.0804
CAI1	1	0.7862	0.2987	6.9271	0.0085
CAI2	2	3.9559	1.5392	6.606	0.0102
CAI3	3	4.2858	1.9387	4.8869	0.0271
RI1		-0.3118	0.1253	6.1893	0.0129
RI2		-1.2637	0.5219	5.863	0.0155
RI3		-1.4334	0.6584	4.7396	0.0295

R-Square	0.0048
Max-rescaled R-Square	0.0124

Table 27

Restatement Magnitude – Robustness Test - Gamma Regression Model

Model 2 – Magnitude of Restatements

Parameter		Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		4.6304	0.0798	3363.88	<.0001
LTA		0.7468	0.0314	564.71	<.0001
OFF		-0.0388	0.0088	19.36	<.0001
NIITI		0.1463	0.0305	23.06	<.0001
OBTA		0.0121	0.015	0.65	0.4209
MU	1	0.7183	0.1311	30.01	<.0001
BHC	1	0.2008	0.0729	7.59	0.0059
CAI1	1	-0.05	0.0594	0.71	0.3999
CAI2	2	0.0583	0.1515	0.15	0.7002
CAI3	3	0.2637	0.1765	2.23	0.1352
Scale		0.5785	0.0135		

Criteria For Assessing Goodness Of Fit

Criterion	Value/DF
Deviance	2.1562

Table 28

Restatement Magnitude – Robustness Test – Heckman Selection Correction Model (Selection Bias)

A. Model 1 – Likelihood of Restatements

Parameter		Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept		-1.4888	0.0324	2117.838	<.0001
LTA		0.0281	0.0133	4.4976	0.0339
OFF		0.0127	0.00385	10.8931	0.001
NIITI		0.0224	0.0104	4.6746	0.0306
OBTA		0.0232	0.00808	8.2017	0.0042
MU	1	0.1722	0.0543	10.065	0.0015
BHC	1	-0.1301	0.0284	20.9912	<.0001
GR		0.00456	0.00808	0.3185	0.5725
LTC		-0.0468	0.0136	11.8628	0.0006
ROAA		-0.0616	0.012	26.4229	<.0001
CAI1	1	0.048	0.0229	4.3933	0.0361
CAI2	2	0.2363	0.0613	14.8769	0.0001
CAI3	3	0.0663	0.0671	0.9757	0.3233

B. Model 2 – Magnitude of Restatements

Parameter		Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		12.371	1.1602	113.7	<.0001
LTA		0.7076	0.0319	492.33	<.0001
OFF		-0.0835	0.0112	55.94	<.0001
NIITI		0.0954	0.031	9.48	0.0021
OBTA		-0.038	0.0178	4.57	0.0326
MU	1	0.0701	0.1629	0.19	0.6667
BHC	1	0.6753	0.1008	44.9	<.0001
CAI1	1	-0.2343	0.0653	12.89	0.0003
CAI2	2	-0.7307	0.1901	14.78	0.0001
CAI3	3	0.0196	0.1793	0.01	0.9131
IMR		-4.0375	0.6021	44.96	<.0001
Scale		0.5858	0.0137		

Criterion	Value/DF
Deviance	2.1263

Next, we look at the conservative recognition of losses and the magnitude of discretionary accruals used in reporting. Based on the significance of RI1 in both **Table 30D** and **Table 32D**, we have an endogeneity issue. Hence, we focus on the results from these models and ignore the results from the Probit Regression **(Table 29)** and Gamma Regression **(Table 31)** models. **Table 30 A, B, C** and **Table 32 A, B, C** presents the results from modeling the decision to have an audit for the four different categories. The residuals from these models are used as independent variables in the modeling of the conservatism (**Table 30D)** and magnitude of discretionary accruals (**Table 32D)**. In our analysis of whether an independent audit increases the propensity for financial reporting conservatism, the results in **Table 30B** indicates that persistently audited banks (CAI1) are less conservative in recognizing probable loan losses. This is consistent with our finding in the main study. For analyzing the effect that an independent audit has on the magnitude of discretionary accruals used in reporting, we find that consistently audited banks (CAI1) had higher magnitudes of discretionary accruals. These results are also consistently with the main study. Interestingly, the results also show that those banks that went from no audit to having an audit (CAI2) had lower magnitudes of discretionary accruals and those banks that went from having an audit to not having an audit had higher magnitudes of discretionary accruals (CAI3). Collectively, the results from these robustness tests confirm the findings in the main study that an independent audit may not improve financial reporting quality and suggest that quality audits are not being performed at small private commercial banks.

Table 29

Discretionary Accruals Conservatism – Robustness Test - Probit Regression Model

Model 4 – Discretionary Accrual Conservatism

Parameter		Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept		-0.1442	0.0103	194.7879	<.0001
SLTL		-0.2724	0.0139	382.7129	<.0001
SΔNPL		-0.1772	0.0161	121.5489	<.0001
SLNPL		-0.2608	0.0102	656.0699	<.0001
SLALL		1.0629	0.0163	4256.644	<.0001
CAI1	1	-0.0377	0.014	7.2691	0.007
CAI2	2	0.00306	0.0463	0.0044	0.9474
CAI3	3	0.0174	0.0465	0.1408	0.7075

Table 30

Discretionary Accruals Conservatism – Robustness Test – Multinomial Probit – Probit Regression Model (Endogeneity)

A. Model 6 – Audit Decision Model (Consistently Audited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	0.2989	0.022	184.1322	<.0001
LTA	0.562	0.0077	5332.011	<.0001
NIITI	0.2043	0.00856	570.0939	<.0001
OBTA	2.4823	0.4261	33.936	<.0001
GR	0.00361	0.00748	0.2329	0.6294
ROAA	-0.1567	0.0108	211.1812	<.0001
MU	1.7186	0.0758	514.3878	<.0001
BHC	-0.2337	0.0202	133.5625	<.0001

B. Model 6 – Audit Decision Model (Unaudited to Audited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	-2.2577	0.0818	761.3198	<.0001
LTA	-0.0806	0.0149	29.2916	<.0001
NIITI	-0.0159	0.0162	0.9565	0.3281
OBTA	-6.7588	2.4569	7.5677	0.0059
GR	0.0345	0.00879	15.3924	<.0001
ROAA	-0.0489	0.0186	6.9308	0.0085
MU	-0.3916	0.1263	9.608	0.0019
BHC	0.0777	0.0422	3.3903	0.0656

C. Model 6 – Audit Decision Model (Audited to Unaudited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	-2.1246	0.0716	879.4655	<.0001
LTA	-0.1267	0.0147	74.2202	<.0001
NIITI	-0.0147	0.016	0.8501	0.3565
OBTA	-3.0475	2.1007	2.1046	0.1469
GR	0.0192	0.00822	5.4495	0.0196
ROAA	-0.0222	0.0187	1.4025	0.2363
MU	-0.3654	0.1223	8.9258	0.0028
BHC	0.0339	0.0409	0.6883	0.4067

D. Model 4 – Discretionary Accrual Conservatism

Parameter		Estimate	StandardError	WaldChi-Square	Pr > ChiSq
Intercept		0.3416	0.123	7.7133	0.0055
SLTL		-0.2453	0.0144	291.4677	<.0001
SΔNPL		-0.1591	0.0162	96.2566	<.0001
SLNPL		-0.25	0.0104	578.7202	<.0001
SLALL		1.0489	0.0165	4041.283	<.0001
CAI1	1	-0.6259	0.0898	48.6056	<.0001
CAI2	2	-2.0556	1.0578	3.776	0.052
CAI3	3	-0.4236	1.3032	0.1057	0.7451
RI1		0.2781	0.0387	51.7032	<.0001
RI2		0.7042	0.358	3.869	0.0492
RI3		0.1555	0.4412	0.1242	0.7245

R-Square	0.2266
Max-rescaled R-Square	0.3039

Table 31

Discretionary Accruals Magnitude – Robustness Test – Multinomial Probit – Gamma Regression Model

Model 5 – Discretionary Accrual Magnitude

Parameter		Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept		-0.6714	0.0069	9595.08	<.0001
SLTL		-0.0521	0.0078	44.41	<.0001
SΔNPL		0.1431	0.0096	224.45	<.0001
SLNPL		0.1148	0.006	366.1	<.0001
SLALL		0.2603	0.0061	1838.86	<.0001
CAI1	1	0.0015	0.0093	0.02	0.8749
CAI2	2	-0.0073	0.0305	0.06	0.8122
CAI3	3	0.0454	0.0306	2.2	0.1376
Scale		1.2954	0.0084		

Criteria For Assessing Goodness Of Fit

Criterion	Value/DF
Deviance	0.8666

Table 32

Discretionary Accruals Magnitude – Robustness Test – Simultaneous Probit - Gamma Regression Model

A. Model 6 – Audit Decision Model (Consistently Audited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	0.2989	0.022	184.1322	<.0001
LTA	0.562	0.0077	5332.011	<.0001
NIITI	0.2043	0.00856	570.0939	<.0001
OBTA	2.4823	0.4261	33.936	<.0001
GR	0.00361	0.00748	0.2329	0.6294
ROAA	-0.1567	0.0108	211.1812	<.0001
MU	1.7186	0.0758	514.3878	<.0001
BHC	-0.2337	0.0202	133.5625	<.0001

B. Model 6 – Audit Decision Model (Unaudited to Audited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	-2.2577	0.0818	761.3198	<.0001
LTA	-0.0806	0.0149	29.2916	<.0001
NIITI	-0.0159	0.0162	0.9565	0.3281
OBTA	-6.7588	2.4569	7.5677	0.0059
GR	0.0345	0.00879	15.3924	<.0001
ROAA	-0.0489	0.0186	6.9308	0.0085
MU	-0.3916	0.1263	9.608	0.0019
BHC	0.0777	0.0422	3.3903	0.0656

C. Model 6 – Audit Decision Model (Audited to Unaudited)

Parameter	Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept	-2.1246	0.0716	879.4655	<.0001
LTA	-0.1267	0.0147	74.2202	<.0001
NIITI	-0.0147	0.016	0.8501	0.3565
OBTA	-3.0475	2.1007	2.1046	0.1469
GR	0.0192	0.00822	5.4495	0.0196
ROAA	-0.0222	0.0187	1.4025	0.2363
MU	-0.3654	0.1223	8.9258	0.0028
BHC	0.0339	0.0409	0.6883	0.4067

D. Model 5 – Discretionary Accrual Magnitude

Parameter		Estimate	Standard Error	WaldChi-Square	Pr > ChiSq
Intercept		-1.1133	0.061	333.26	<.0001
SLTL		-0.0631	0.008	62.2	<.0001
SΔNPL		0.1373	0.0096	204.02	<.0001
SLNPL		0.1091	0.0061	320.38	<.0001
SLALL		0.2737	0.0063	1901.99	<.0001
CAI1	1	0.4684	0.0463	102.18	<.0001
CAI2	2	-1.6635	0.5618	8.77	0.0031
CAI3	3	7.3312	0.6799	116.26	<.0001
RI1		0.5793	0.1071	29.27	<.0001
RI2		1.2021	0.172	48.83	<.0001
RI3		-1.8236	0.2496	53.4	<.0001
Scale		1.3013	0.0084		

Criteria For Assessing Goodness Of Fit

Criterion	Value/DF
Deviance	0.8625

Table 33

Summary of Results

	Hypotheses	Results
Hypothesis 1	The decision to have an independent audit is a systematic and endogenous.	Based on the high accuracy rates of the Simple Logistic, JRIP, Random Forest, and SMO machine learning algorithms, this hypothesis was supported and indicates that the decision to have an independent audit may be systematic and is endogenously and optimally determined given company characteristics.
Hypothesis 2	An independent audit decreases the likelihood of having restatement.	The hypothesis was not supported and the results show that an independent audit may not decrease the likelihood of a restatement. Independently audited banks were found to be more likely to have a restatement then unaudited banks.
Hypothesis 3	An independent audit decreases the magnitudes of restatements.	The hypothesis was supported and indicates that an independent audit may reduce the magnitude of restatements. Collectively, Hypothesis 2 and Hypothesis 3 may suggest that independently audited banks had a higher likelihood of restatements but the magnitude of those restatements were smaller than unaudited banks.
Hypothesis 4	An independent audit increases the conservative recognition of expected and probable losses.	The hypothesis was not supported. The results show that independently audited banks were less conservative than unaudited banks in the recognition of probable loan losses.

Hypothesis 5	An independent audit decreases the magnitudes of discretionary accruals in financial reporting.	The hypothesis was not supported. The findings indicate that independently audited banks had higher magnitudes of discretionary accruals than unaudited banks. Collectively, Hypothesis 4 and Hypothesis 5 may suggest that independently audited banks were more aggressive in recognizing income by under reporting probable loan losses with greater magnitude.

Chapter 6

6. LIMITATIONS AND CONCLUDING REMARKS

We find that machine learning algorithms are able to predict with reasonable accuracy whether a bank is audited or not audited utilizing only bank characteristic variables. The Simple Logistic, Support Vector Machine (SMO), JRIP, and Random Forest machine learning algorithms appear to have the best potential to accurately classify whether private commercial banks in the dataset are audited or not audited. 10-fold cross-validation was used in the evaluation to show that the results are robust and do not suffer from over fitting. The high accuracy rates (above 70%) of these algorithms indicate that the decision to have an audit is systematic and endogenous. However, a limitation must be kept in mind that these results are specific to the FDIC commercial bank dataset used and may not be generalizable (perform with similar accuracy in future observations or other private companies). We also find specifically that profitability, growth, complex operations, and hierarchal or ownership structure can influence the decision to have an audit. In future studies, the Simple Logistic, Support Vector Machine (SMO), JRIP, and Random Forest machine learning algorithms can be tested on other private company datasets to determine the generalizability of our results.

The main contribution of this dissertation to the accounting and auditing literature is to provide insight or direct evidence on whether an independent audit increases financial reporting quality in the small private commercial bank setting. We use material accuracy, conservative recognition of losses, and the magnitude of discretionary accruals used in reporting to measure financial reporting quality. The study first finds that audited banks had a greater probability of having a restatement. However, these restatements were of lower magnitude than restatements from unaudited banks. Second, we find that an independent audit does not promote conservatism in financial reporting. More specifically, the results from the study indicate that audited banks are less conservative in recognizing probable loan losses. Furthermore, audited banks had higher magnitudes of discretionary accruals in reporting than unaudited banks. Collectively, these findings provide evidence that an independent audit may not increase the quality of financial reporting and question the value of an independent audit absent regulatory requirements. Furthermore, the findings provide implications on the quality of audits being performed at small private commercial banks. Future researchers may want to examine other industries that do not have regulatory audit requirements and provide insights on whether the findings in this study hold true for other private company audits. Additionally, future researchers may want to use other measurements of financial reporting quality to see if an independent audit improves those measurements.

The results of this study are bewildering and may be hard to accept. However, our results are supported and substantiated by rigorous analysis and robustness tests. On the other hand, it can be argued that the results from the study are not surprising based on recent popular press and research. (Whalen and Cheffers 2012) find that

21.8% of Russell 1000 companies audited by the Big Four had errors in their financial statements and required restatements. A former SEC Chief Accountant, Lynn Turner, comments that the results from (Whalen and Cheffers 2012) study call into question the quality of audits, competence of the CFO/Controllers, quality of internal controls, and the role of the audit committee in financial reporting quality. Other recent studies (Analytics 2007; Plumlee and Yohn 2009; Scholz 2008; Taub 2006; Turner and Weirich 2006) recognized that restatements by public companies are on the rise. Collectively, the results from these recent studies and the findings from this study suggest that regulators and researchers should look further into the relationship between an independent audit and financial reporting quality. Furthermore, audit firms should look at their audit methodologies and perhaps innovate to do better audits. Audit firms may want to consider the use of audit automation such continuous auditing and monitoring to innovate the audit process. The use of advanced audit technology such as continuous auditing and monitoring may enhance the traditional audit process by improving its efficiency and effectiveness (Chan and Vasarhelyi 2010). The concept of continuous auditing was first introduced by (Groomer and Murthy 1989) and (Vasarhelyi and Halper 1991). In the continuous auditing environment, the whole population of economic transactions is considered by using automated analytical monitoring (Vasarhelyi et al. 2004; Kogan et al. 2010). The consideration of the whole population provides a more comprehensive or effective audit to detect material misstatements.

The study has a number of limitations. First, we constrain ourselves to one industry, and hence the results may not be generalizable. However, there is an advantage of studying small private commercial banks. Small private commercial banks with under $500 million in total assets are not required to have an independent audit. The voluntary decision to have an audit by a homogenous large population allows us to analyze and isolate whether an independent audit has any propensity to increase financial reporting quality. Second, we do not have data on which auditors audited the banks. In the literature, audits by the Big 4 are synonymous with quality audits and quality financial reporting. Third, the Call Reports are not specifically audited by the auditors but the data extracted from the general ledger (GL) which are used to generate the regulatory reports are audited. Fourth, an audited bank without a restatement does not directly indicate quality auditing. The bank being audited could have had excellent controls over financial reporting and produced high quality financial reports without any regards to audit quality. Fifth and last, we estimate the non-discretionary and discretionary component of the allowance for loan losses using statistical models without context of the management's and the auditor's judgment of the loan portfolio quality. Perhaps there are unobserved justifiable reasons to inflate or deflate the quality of loans by management and their auditors. We also make assumptions that our models accurately model the nondiscretionary and discretionary component of the allowance for loan losses. These assumptions are necessary and we understand they may affect our findings.

REFERENCES

Abbott, L. J., Y. Park, and S. Parker. 2000. The effects of audit committee activity and independence on corporate fraud. *Managerial Finance* 26 (11):55-68.

Abdel-Khalik, A. R. 1993. Why Do Private Companies Demand Auditing? A Case for Organizational Loss of Control. *Journal of Accounting, Auditing & Finance* 8 (1):31-52.

Aghaie, A., and A. Saeedi. 2009. Using Bayesian Networks for Bankruptcy Prediction: Empirical Evidence from Iranian Companies. In *2009 International Conference on Information Management and Engineering*.

Aha, D. W., D. Kibler, and M. K. Albert. 1991. Instance-based learning algorithms. *Machine Learning* 6 (1):37-66.

AICPA. 2004. Depository and lending institutions : banks and savings institutions, credit unions, finance companies and mortgage companies. *Depository and lending institutions : banks and savings institutions, credit unions, finance companies and mortgage companies.*

Aier, J. K., J. Comprix, M. T. Gunlock, and D. Lee. 2005. The Financial Expertise of CFOs and Accounting Restatements. *Accounting Horizons* 19 (3):123-135.

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

Altman, E. I., R. G. Haldeman, and P. Narayanan. 1977. ZETATM analysis A new model to identify bankruptcy risk of corporations. *Journal of Banking & Finance* 1 (1):29-54.

Analytics, A. 2007. Financial restatements and market reactions: Ives Group Inc.

Ball, R., and L. Shivakumar. 2005. Earnings quality in UK private firms: comparative loss recognition timeliness. *Journal of Accounting and Economics* 39 (1):83-128.

Bartov, E., F. A. Gul, and J. S. L. J. S. L. Tsui. 2000. Discretionary-accruals models and audit qualifications. *Journal of Accounting and Economics* 30 (3):421-452.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings1. *Journal of Accounting and Economics* 24 (1):3-37.

Basu, S., L. Hwang, and C.-L. Jan. 2002. Differences in Conservatism between Big Eight and non-Big Eight Auditors. In *Working Paper*: Baruch College, City University of New York.

Beasley, M. S. 1996. An Empirical Analysis of the Relation Between the Board of Director Composition and Financial Statement Fraud. *Accounting Review* 71 (4):443-465.

Beatty, A., and D. G. Harris. 1999. The Effects of Taxes, Agency Costs and Information Asymmetry on Earnings Management: A Comparison of Public and Private Firms. *Review of Accounting Studies* 4 (3):299-326.

Beatty, A. L., B. Ke, and K. R. Petroni. 2002. Earnings Management to Avoid Earnings Declines across Publicly and Privately Held Banks. *The Accounting Review* 77 (3):547-570.

Beaver, W. H. 1966. Financial Ratios As Predictors of Failure. *Journal of Accounting Research* 4:71-111.

Beaver, W. H., and E. E. Engel. 1996. Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting and Economics* 22 (1-3):177-206.

Becker, C. L., M. L. Defond, J. Jiambalvo, and K. R. Subramanyam. 1998. The Effect of Audit Quality on Earnings Management*. *Contemporary Accounting Research* 15 (1):1-24.

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

Bell, T. B., and R. H. Tabor. 1991. Empirical Analysis of Audit Uncertainty Qualifications. *Journal of Accounting Research* 29 (2):350-370.

Breiman, L. 2001. Random Forests. *Machine Learning* 45 (1):5-32.

Cano-Rodríguez, M. 2010. Big Auditors, Private Firms and Accounting Conservatism: Spanish Evidence. *European Accounting Review* 19 (1):131-159.

Chan, D. Y., and M. A. Vasarhelyi. 2010. Innovation and Practice of Continuous Auditing. *International Journal of Accounting Information Systems* Forthcoming.

———. 2011. Innovation and practice of continuous auditing. *International Journal of Accounting Information Systems* 12 (2):152-160.

Chen, K. C. W., and B. K. Church. 1992. Default on Debt Obligations and the Issuance of Going-Concern Opinions. *AUDITING* 11 (2):30-49.

Chow, C. W. 1982. The Demand for External Auditing: Size, Debt and Ownership Influences. *Accounting Review* 57 (2):272.

Chung, R., M. Firth, and K. Jeong-Bon. 2003. Auditor conservatism and reported earnings. *Accounting & Business Research* 33 (1):19-32.

Cohen, J., G. Krisnamoorthy, and A. M. Wright. 2002. Corporate Governance and the Audit Process. *Contemporary Accounting Research* 19 (4):573-594.

Cohen, J. R., G. Krishnamoorthy, and A. Wright. 2004. The Corporate Governance Mosaic and Financial Reporting Quality. *Journal of Accounting Literature*:87-152.

Collis, J. 2010. Audit Exemption and the Demand for Voluntary Audit: A Comparative Study of the UK and Denmark. *International Journal of Auditing* 14 (2):211-231.

Collis, J., R. Jarvis, and L. Skerratt. 2004. The demand for the audit in small companies in the UK. *Accounting & Business Research* 34 (2):87-100.

Cook, R. D. 1977. Detection of Influential Observation in Linear Regression. *Technometrics* 19 (1):pp. 15-18

Dahl, D., J. P. O'Keefe, and G. A. Hanweck. 1998. The Influence of Examiners and Auditors on Loan-Loss Recognition. *FDIC Banking Review*.

DeAngelo, L. E. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (3):183-199.

DeFond, M. L., and K. R. Subramanyam. 1998. Auditor changes and discretionary accruals. *Journal of Accounting and Economics* 25 (1):35-67.

Dopuch, N., R. W. Holthausen, and R. W. Leftwich. 1987. Predicting Audit Qualifications with Financial and Market Variables. *Accounting Review* 62 (3):431.

Fanning, K. M., and K. O. Cogger. 1998. Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems in Accounting, Finance & Management* 7:21-41.

Flint, D. 1988. *Philosophy and principles of auditing : an introduction*. Basingstoke: Macmillan Education.

Francis, J. R., and J. Krishnan. 1999. Accounting Accruals and Auditor Reporting Conservatism*. *Contemporary Accounting Research* 16 (1):135-165.

Francis, J. R., and D. Wang. 2004. Investor Protection, Auditor Conservatism and Earnings Quality: Are Big 4 Auditors Conservative only in the United States? In *Working Paper*.

Francis, J. R., D. Wang, and A. Nikitkov. 2003. The Effect of Legal Environment on Big Five Auditor Conservatism Around the World. In *Working Paper*.

Green, B. P., and J. H. Choi. 1997. Assessing the Risk of Management Fraud Through Neural Network Technology. *AUDITING* 16 (1):14-28.

Greene, W. H. 2003. *Econometric analysis*. Upper Saddle River, N.J.: Prentice Hall.

Groomer, S. M., and U. S. Murthy. 1989. Continuous Auditing of Database Applications: An Embedded Audit Module Approach. *Journal of Information Systems* 3 (2):53.

Gunther, J. W., and R. R. Moore. 2003. Loss underreporting and the auditing role of bank exams. *Journal of Financial Intermediation* 12 (2):153-177.

Hansen, J. V., J. B. McDonald, W. F. Messier, Jr., and T. B. Bell. 1996. A Generalized Qualitative-Response Model and the Analysis of Management Fraud. *Management Science* 42 (7):1022-1032.

Hastie, T., and R. Tibshirani. 1998. Classification by Pairwise Coupling. *The Annals of Statistics* 26 (2):451-471.

Heckman, J. J. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47 (1):153-161.

Heninger, W. G. 2001. The Association between Auditor Litigation and Abnormal Accruals. *The Accounting Review* 76 (1):111-126.

Imai, K., G. King, and O. Lau. 2007a. gamma: Gamma Regression for Continuous, Positive Dependent Variables. In *Zelig: Everyone's Statistical Software*.

———. 2007b. probit: Probit Regression for Dichotomous Dependent Variables. In *Zelig: Everyone's Statistical Software*.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (4):305-360.

John, G. H., and P. Langley. 1995. Estimating continuous distributions in Bayesian classifiers. Paper read at Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, at Montreal, Quebec.

Johnson, E., I. K. Khurana, and J. K. Reynolds. 2002. Audit-Firm Tenure and the Quality of Financial Reports*. *Contemporary Accounting Research* 19 (4):637-660.

Keasey, K., R. Watson, and P. Wynarczyk. 1988. The Small Company Audit Qualification: A Preliminary Investigation. *Accounting & Business Research* 18 (72):323-333.

Keerthi, S. S., S. K. Shevade, C. Bhattacharyya, and K. R. K. Murthy. 2001. Improvements to Platt's SMO Algorithm for SVM Classifier Design. *Neural Computation* 13 (3):637-649.

Kim, Y., D. Y. Chan, A. Kogan, and M. A. Vasarhelyi. 2011. A Model to Detect Potentially Fraudulent/Abnormal Wires of an Insurance Company: An Unsupervised Rule-based. In *Working Paper*.

Kinney Jr, W. R., and R. D. Martin. 1994. Does Auditing Reduce Bias in Financial Reporting? A Review of Audit-Related Adjustment Studies. *AUDITING* 13 (1):149-156.

Kinney, W. R., and L. S. McDaniel. 1989. Characteristics of firms correcting previously reported quarterly earnings. *Journal of Accounting and Economics* 11 (1):71-93.

Kirkos, E., C. Spathis, and Y. Manolopoulos. 2007. Data Mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications* 32 (4):995-1003.

Kogan, A., M. A. Vasarhelyi, and J. Wu. 2010. Continuous Data Level Auditing Using Continuity Equations. *Working paper, Rutgers Business School*.

Kohlbeck, M. J. 2005. The Demand for Private Company Audits: Evidence from Private Commercial Banks. *SSRN eLibrary*.

Kreutzfeldt, R., and W. Wallace. 1987. Error characteristics in audit populations: their profile and relationship to environmental factors. *Auditing: A Journal of Practice & Theory* 5:20-43.

Lazer, R., J. Livnat, and C. E. L. Tan. 2004. Restatements and Accruals after Auditor Changes. In *Working Paper*: New York University.

Lee, P. J., S. J. Taylor, and S. L. Taylor. 2006. Auditor Conservatism and Audit Quality: Evidence from IPO Earnings Forecasts. *International Journal of Auditing* 10 (3):183-199.

Levitan, A. S., and J. A. Knoblett. 1985. Indicators of Exceptions to the Going Concern Assumption. *AUDITING* 5 (1):26.

Lindo, A. 2007. Financial Restatements: Observations from the Federal Reserve. Paper read at AICPA National Banks & Savings Institutions Conference.

Lobo, G. J., and J. Zhou. 2006. Did Conservatism in Financial Reporting Increase after the Sarbanes-Oxley Act? Initial Evidence. *Accounting Horizons* 20 (1):57-73.

Martens, D., L. Bruynseels, B. Baesens, M. Willekens, and J. Vanthienen. 2008. Predicting going concern opinion with data mining. *Decision Support Systems* 45 (4):765-777.

McKee, T. E. 1998. A mathematically derived rough set model for bankruptcy prediction. In *In C. E. Brown (Ed.), Collected papers of the seventh annual research workshop on artificial intelligence and emerging technologies in accounting, auditing and tax, artificial intelligence/emerging technologies section of the american accounting association.*

McKee, T. E., and M. Greenstein. 2000. Predicting bankruptcy using recursive partitioning and a realistically proportioned data set. *Journal of Forecasting* 19 (3):219-230.

McKee, T. E., and T. Lensberg. 2002. Genetic programming and rough sets: A hybrid approach to bankruptcy classification. *European Journal of Operational Research* 138 (2):436-451.

Menon, K., and K. B. Schwartz. 1987. An empirical investigation of audit qualification decisions in the presence of going concern uncertainties*. *Contemporary Accounting Research* 3 (2):302-315.

Mutchler, J. F. 1985. A Multivariate Analysis of the Auditor's Going-Concern Opinion Decision. *Journal of Accounting Research* 23 (2):668-682.

Odom, M. D., and R. Sharda. 1990. A neural network model for bankruptcy prediction. Paper read at Neural Networks, 1990., 1990 IJCNN International Joint Conference on.

Ohlson, J. A. 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research* 18 (1):109-131.

Park, C.-S., and I. Han. 2002. A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. *Expert Systems with Applications* 23 (3):255-264.

Pierre, K. S., and J. A. Anderson. 1984. An Analysis of the Factors Associated with Lawsuits against Public Accountants. *The Accounting Review* 59 (2):242-263.

Platt, J. C. 1999. Fast training of support vector machines using sequential minimal optimization. In *Advances in kernel methods*: MIT Press, 185-208.

Plumlee, M., and T. L. Yohn. 2010. An Analysis of the Underlying Causes Attributed to Restatements. *Accounting Horizons* 24 (1):41-64.

Plumlee, M. A., and T. L. Yohn. 2009. An Analysis of the Underlying Causes Attributed to Restatements. In *Working Paper*: University of Utah.

Poirier, D. J. 1980. Partial observability in bivariate probit models. *Journal of Econometrics* 12 (2):209-217.

Rennie, M., D. Senkow, R. Rennie, and J. Wong. 2003. DEREGULATION OF THE PRIVATE CORPORATION AUDIT IN CANADA: JUSTIFICATION, LOBBYING, AND OUTCOMES. *Research in Accounting Regulation* 16:227-241.

Sarkar, S., and R. S. Sriram. 2001. Bayesian Models for Early Warning of Bank Failures. *Management Science* 47 (11):1457-1475.

Scholz, S. 2008. The Changing Nature and Consequences of Public Company Financial Restatements: 1997-2006: The Department of the Treasury.

Seow, J.-L. 2001. The Demand for the UK Small Company Audit an Agency Perspective. *International Small Business Journal* 19 (2):61-79.

Shin, K.-S., T. S. Lee, and H.-j. Kim. 2005. An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications* 28 (1):127-135.

Shumway, T. 2001. Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business* 74 (1):101-124.

Solomon, I., M. D. Shields, and O. R. Whittington. 1999. What Do Industry-Specialist Auditors Know? *Journal of Accounting Research* 37 (1):191-208.

Spathis, C., M. Doumpos, and C. Zopounidis. 2002. Detecting falsified financial statements: a comparative study using multicriteria analysis and multivariate statistical techniques. *European Accounting Review* 11 (3):509-535.

Staub, K. E. 2009. Simple Tests for Exogeneity of a Binary Explanatory Variable in Count Data Regression Models. *Communications in Statistics: Simulation & Computation* 38 (9):1834-1855.

Summers, S. L., and J. T. Sweeney. 1998. Fraudulently Misstated Financial Statements and Insider Trading: An Empirical Analysis. *Accounting Review* 73 (1):131.

Sumner, M., E. Frank, and M. Hall. 2005. Speeding Up Logistic Model Tree Induction. In *Knowledge Discovery in Databases: PKDD 2005*, edited by A. Jorge, L. Torgo, P. Brazdil, R. Camacho and J. Gama: Springer Berlin / Heidelberg, 675-683.

Sun, L., and P. P. Shenoy. 2007. Using Bayesian networks for bankruptcy prediction: Some methodological issues. *European Journal of Operational Research* 180 (2):738-753.

Tam, K. Y., and M. Y. Kiang. 1992. Managerial Applications of Neural Networks: The Case of Bank Failure Predictions. *Management Science* 38 (7):926-947.

Tan, P.-N., M. Steinbach, and V. Kumar. 2005. *Introduction to data mining*. Boston: Pearson Addison Wesley.

Taub, S. A. 2006. Speech by SEC Staff: Remarks Regarding Restatements Before the Financial Executives International Meeting, edited by U. S. S. a. E. Commission.

Tauringana, V., and S. Clarke. 2000. The demand for external auditing: managerial share ownership, size, gearing and liquidity influences. *Managerial Auditing Journal* 15 (4):160-168.

Thiprungsri, S. 2011. Cluster Analysis for Anomaly Detection in Accounting Data. In *Working Paper*.

Thomaidis, N., G. Gounias, and C. Zopounidis. 1999. A fuzzy rule based learning method for corporate bankruptcy prediction. Paper read at ACAI 99, at Chania, Greece.

Turner, J. L., and J. T. Sennetti. 2001. Post-Audit Restatement Risk and Audit Firm Size. *Journal of Forensic Accounting* 2:67--94.

Turner, L. E., and T. R. Weirich. 2006. A Closer Look at Financial Statement Restatements: Analyzing the Reasons Behind the Trend. *The CPA Journal*.

Van Tendeloo, B., and A. Vanstraelen. 2008. Earnings Management and Audit Quality in Europe: Evidence from the Private Client Segment Market. *European Accounting Review* 17 (3):447-469.

Vasarhelyi, M. A., M. G. Alles, and A. Kogan. 2004. Principles of Analytic Monitoring for Continuous Assurance. *Journal of Emerging Technologies in Accounting* 1 (1):1-21.

Vasarhelyi, M. A., and F. B. Halper. 1991. The Continuous Audit of Online Systems. *Auditing: A Journal of Practice & Theory* 10 (1).

Wahlen, J. M. 1994. The Nature of Information in Commercial Bank Loan Loss Disclosures. *The Accounting Review* 69 (3):455-478.

Wallace, W. A. 1980. *The economic role of the audit in free and regulated markets*. S.l.: s.n.].

Watts, R. L., and J. L. Zimmerman. 1983. Agency Problems, Auditing, and the Theory of the Firm: Some Evidence. *Journal of Law & Economics* 26 (1):613-634.

Whalen, D., and M. Cheffers. 2012. A Restatement Analysis of the Russell 1000 Companies: The Extent to which the "Fresh Eyes" of a Newly Engaged Auditor Provided Assistance in the Discover of the Misstatement: Audit Analytics.

Witten, I. H., and E. Frank. 2005. *Data mining : practical machine learning tools and techniques*. Amsterdam; Boston, MA: Morgan Kaufman.

Wooldrige, J. M. 1997. Quasilikelihood methods for count data. In *Handbook of Applied Econometrics*. Massachusets, USA/Oxford, UK: Blackwell Publishers, 352-406.

Zhang, G., M. Y. Hu, B. Eddy Patuwo, and D. C. Indro. 1999. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research* 116 (1):16-32.

Zhou, J. 2007. Financial Reporting After the Sarbanes-Oxley Act: Conservative or Less Earnings Management? *Research in Accounting Regulations, Vol. 20, pp. 187-192*.

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

APPENDIX

Variable Definitions

RS	=	Restatement (Dummy Variable - 1 if the bank restated and 0 otherwise);
RSABS	=	Absolute value of Restatement;
LTA	=	Log of Total Assets;
OFF	=	Number of offices, branches, locations, and facilities;
NIITI	=	Non-interest Income divided by total Interest and Non-interest income;
OBTA	=	Off-balance sheet activities divided by total assets;
MU	=	Mutual or stockholder bank (Dummy Variable - 1 if a mutual bank and 0 a stock bank);
BHC	=	Parent is a Bank Holding Company (Dummy Variable - 1 if the bank's parent is a bank holding company and 0 otherwise);
GR	=	Change in total assets divided by beginning total assets;
LTC	=	Total loans divided by Total Equity Capital;
ROAA	=	Net income (Loss) divided by Total Average Assets (assets at the end of the previous year plus assets at the end of the current year divided by 2);
DALL	=	Allowance for loan losses;
DCO	=	Loan charge offs;
DTL	=	Total Loans;
DNPL	=	Nonperforming loans;

ΔDNPL	=	Change in nonperforming loans as a percentage of the average of beginning and ending total loans;
rALLDirect	=	Discretionary component of allowance for loan losses (Residual from (Beaver and Engel 1996) Model);
rALLABS	=	Magnitude of discretionary component of allowance for loan losses (Absolute value of residual from (Beaver and Engel 1996) Model;
SLTL	=	Beginning balance of total loans;
SΔNPL	=	Change in Nonperforming Loans;
SLNPL	=	Beginning balance of nonperforming loans;
SLALL	=	Beginning balance of allowance for loan losses;
AI	=	Audit Indicator (Dummy Variable - 1 if the bank is independently audited and 0 otherwise);
CAI1	=	Audit Indicator (Dummy Variable – 1 if the bank is independently audited in the prior period and independently audited in the current period and 0 otherwise);
CAI2	=	Audit Indicator (Dummy Variable – 1 if the bank is not Independently audited in the prior period and independently audited in the current period and 0 otherwise); and
CAI3	=	Audit Indicator (Dummy Variable – 1 if the bank is independently audited in the prior period and not independently audited in the current period and 0 otherwise).

CURRICULUM VITA

David Y. Chan

1979	Born in New York, NY USA
1993-97	High School for Health Professions & Human Services, Regents Diploma
1997-01	St. John's University, B.S. in Finance
2002-04	St. John's University, M.S. in Accounting
2004-05	Grant Thornton LLP, Assurance Associate
2006-08	Self Employment, Certified Public Accountant
2007	Queensborough Community College, Lecturer
2007-08	St. John's University, Lecturer
2008-12	Rutgers University, Ph.D. in Accounting and Information Systems