ANALYTICS WITH EXCEPTION PRIORITIZATION, CONSUMER SEARCH VOLUME, AND SOCIAL CAPITAL

By PEI LI

A dissertation submitted to the

Graduate School- Newark

Rutgers, The State University of New Jersey

in partial fulfillment of the 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. Kevin Moffitt

____

Dr. Graham Gal

Newark, New Jersey

May, 2016

ABSTRACT OF THE DISSERTATION

Analytics with Exception Prioritization, Consumer Search Volume, and Social Capital

by PEI LI

Dissertation Director: Dr. Alexander Kogan

This dissertation comprises three essays. The first essay addresses the issue of the large volume of exceptions generated by continuous auditing systems. A framework that uses the theory of belief functions is proposed to systematically prioritize exceptions based on the likelihood of an exception being erroneous. The evaluation of the proposed framework is implemented using a simulated experiment. The results of the experiment indicate that the framework has the potential to effectively prioritize exceptions.

The second essay examines whether the consumer search volume can be employed as a type of nonfinancial information in analytical procedures to improve the accuracy of prediction and error detection. This study finds that the model that incorporates the consumer search volume generally outperforms the benchmark models in terms of prediction and error detection in analytical procedures.

The third essay examines the impact of social capital on the municipal bond market. The municipalities with high social capital are expected to be more trustworthy and likely to honor their debt obligations. The results show that municipalities located in the high social capital areas issue bonds with lower yields. The findings from the

secondary market also show that bonds issued by the municipalities located in the high social capital areas have higher prices.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the help of dissertation committee members and the professors that I have worked with.

Above all, I would like to give a special thanks to my advisor Dr. Alexander Kogan for your continuous trust, support, and patience. I learn how to be a serious researcher from you. You have been always there to answer my questions and provide suggestions.

I would like to express my gratitude to Dr. Miklos Vasarhelyi for your advice, caring and patience. Your academic spirit will always inspire me.

I would like to express my gratitude to Dr. Kevin Moffitt and Dr. Graham Gal for your advice and helpful suggestions.

I am grateful to Dr. Bikki Jaggi for your help and valuable advice on research. I would also like to thank Dr. David Chan for helping me with the first chapter of this dissertation. I thank Dr. Michael Alles and Dr. Valentin Dimitrov for teaching me to have a critical mindset for research.

Finally, I would like to thank my family for your endless love and support throughout my doctoral studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Audit analytics has been applied to improve the efficiency and effectiveness of auditing. The application of information technologies such as continuous auditing has propelled the auditing practice into the modern era. The application of analytical methods gives auditors the opportunity to systematically analyze corporate performance. The first part of this dissertation consists of two essays that utilize analytical methods to improve the efficiency and effectiveness of auditing.

The second part of this dissertation examines how social capital influences the municipal bond market. Economists and sociologists introduced social capital to represent the collective value of social networks or groups. They argue that social capital facilitates cooperation and fosters greater trust over time among people (Woolcock 2001; Guiso et al 2004). Social capital is believed to enhance the performance of local governments, improve economic development, and reduce the transaction costs (Putnam 1995; Fukuyama 1995). The third essay exploits unique features of the municipal bond market to assess how social capital affects the municipal bond market.

The first essay is presented in Chapter 1. Researchers have found that the volume of exceptions generated by a continuous auditing system can be overwhelming for an internal audit department to handle (Alles, Brennan, Kogan, and Vasarhelyi 2006; Alles, Kogan, and Vasarhelyi 2008; and Perols and Murthy 2012). This essay proposes and validates a framework that systematically prioritizes exceptions based on the likelihood of an exception being erroneous. The framework consists of six stages: 1) generation of exceptions using defined rules, 2) assignment of suspicion scores to exceptions using belief functions, 3) exception prioritization, 4) exception investigation, 5) rule confidence level update utilizing back propagation, and 6) rule(s) addition utilizing a rule learner algorithm. This essay evaluates the proposed framework using a simulation experiment. The experiment results provide evidence that the framework can be effective in prioritizing exceptions and thus maximize audit efficiency.

Chapter 2 of the dissertation examines the use of consumer search volume as a source of nonfinancial information used in analytical procedures. This essay examines whether the consumer search volume can be employed as nonfinancial information in analytical procedures to improve the accuracy of prediction and help detect errors in the financial statements. The consumer search volume captures the general level of consumer interest for corporate products or services. The consumer search volume is measured using the search volume index reported by Google. This study finds that the model incorporating the consumer search volume generally outperforms the benchmark models when generating prediction in analytical procedures. This study further examines the error detection ability of the consumer search volume through a simulated experiment. The results show that the model incorporating the consumer search volume generates smaller false positives and false negatives relative to the benchmark model.

Chapter 3 of this dissertation examines the impact of social capital on the debt costs in the municipal bond market. The municipal bond market is subject to a much lesser degree of the Security Exchange Commission's (SEC) reporting requirements compared to the corporate capital markets. In view of the limited regulatory oversight in this market, higher environmental social norms are likely to encourage the municipal officials to have greater transparency in financial information, which will reduce information asymmetry between bond issuers and bondholders (Styles and Tennyson, 2007). This essay argues that high social capital of counties in which municipalities are located enhances the trustworthiness of information on municipal bonds and thus reduces the debt costs. The findings show that bonds issued by municipalities located in high social capital counties exhibit lower cost of debt, reflected by a lower bond yield, compared to the municipalities located in low social capital counties. This finding supports the expectation that the trustworthiness generated by high social capital lowers market risk for municipal bonds because of the higher credibility of bond issuers that enhances the reliability of bond information. The findings are also supported by bond prices in the secondary market. These findings are, however, valid only for general obligation bonds and the social capital seems to have no impact on revenue bonds. Additionally, the results show that bonds are less likely to have insurance when issued by municipalities in high social capital counties relative to municipalities in low social capital counties.

The remainder of the dissertation is organized as follows. Chapter 1 proposes and validates a framework that systematically prioritizes exceptions based on the likelihood of an exception being erroneous and evaluates the framework using an experiment. Chapter 2 examines the contribution of the consumer search volume to analytical procedures. Chapter 3 examines the impact of social capital on the cost of debt in the municipal bond market.

Chapter 1 Exception Prioritization In Continuous Auditing: A Framework And Experimental Evaluation

1.1. Introduction

The concept of continuous auditing was first introduced by Groomer and Murthy (1989) and Vasarhelyi and Halper (1991). A continuous audit involves the use of computers to automate audit procedures, thereby enabling real time or near real time assurance. In theory, a continuous audit should increase efficiency and therefore reduce the cost of auditing. However, the inherent nature of a continuous auditing (CA) system may in fact diminish any economic benefits from automation. Researchers find that a CA system can generate a large volume of exceptions (Alles, Brennan, Kogan, and Vasarhelyi 2006; Alles, Kogan, and Vasarhelyi 2008; and Perols and Murthy 2012). Exceptions are irregular or suspicious transactions, or internal controls violations identified by the CA system. These exceptions need to be manually investigated by auditors. Consequently, a large number of exceptions may become overwhelming and thus limit or negate any economic efficiency gains through automation.

Continuous audit systems are primarily implemented and maintained by internal auditors (Kuenkaikaew and Vasarhelyi 2013; Byrnes et al. 2015a; Byrnes et al. 2015b). An internal audit function can be expensive to operate every year. As a result, management and their internal audit department have the incentive to maximize their audit resources when investigating exceptions generated by a CA system. This study proposes and validates a framework that prioritizes exceptions and simulates the framework using an experiment. The framework consists of six stages: 1) generation of exceptions using defined rules, 2) assignment of suspicion scores to exceptions using belief functions, 3) exception prioritization, 4) exception investigation, 5) rule confidence level update utilizing back propagation, and 6) rule(s) addition utilizing a rule learner algorithm.

The purpose of the framework is to develop a methodology that prioritizes exceptions and directs auditors to focus their resources on investigating highly suspicious exceptions. The exception prioritization framework is centered on an initial set of rules that are defined by internal auditors to detect irregular transactions. Irregular transactions are named as errors in this study for the remaining text. These rules are assigned a confidence level depending on their potency in detecting errors. The CA system identifies transactions that violate a single rule or multiple rules and labels those transactions as exceptions. The suspicion of each of these exceptions is determined using the Dempster-Shafer theory of belief functions and the auditors will be directed to investigate those exceptions with the highest suspicion scores.

The framework has an advanced feature that learns from positively identified errors after each iterative process. Based on the investigative results, the confidence level of a rule that contributed to finding erroneous transactions is revised accordingly. Transactions that violate these rules going forward will receive a higher suspicion score and therefore have a higher priority for investigation. This mechanism is commonly referred to in the machine learning literature as back propagation. Lastly, a rule learner algorithm is implemented to add new rules to the original set of rules that were developed by auditors. This feature captures attributes of positively identified cases of errors to create new rules that will attempt to find similar instances subsequently.

For validation purposes, the framework is evaluated in this study using an experiment. The framework's performance is evaluated by: 1) the ability to effectively prioritize erroneous transactions higher than regular transactions; and 2) the ability to improve the prioritization performance after each iterative process. The results from the experiment provide evidence that the proposed framework has the ability to effectively prioritize true erroneous transactions. Furthermore, the results indicates that using back propagation to refine the confidence levels of rules and using a rule learner algorithm to generate additional rules helped improve the effectiveness of the prioritization of exceptions in subsequent iterative processes.

The remainder of the paper is organized as follows. Section 1.2 reviews the literature on continuous auditing and the theoretical development and application of belief functions in the auditing domain. Section 1.3 discusses the exception prioritization framework in detail. Section 1.4 discusses the data and the rules that are used in this study. Section 1.5 elaborates on the experiment and evaluates the framework. Section 1.6 presents the results from the experiment, while the robustness testing of the framework is discussed in Section 1.7. Lastly, Section 1.8 concludes the paper and presents future research directions.

1.2. Literature Review

Continuous Auditing

Continuous auditing is a method of auditing that produces audit results simultaneously with or a short period of time after the occurrence of relevant events (Kogan, Sudit, and Vasarhelyi 1999). In the continuous auditing environment, audit procedures are automated using computers to enable real time or near real time assurance. The concept of continuous auditing was first introduced over two and half decades ago by Groomer and Murthy (1989) and Vasarhelyi and Halper (1991). However, the practical innovation and application of continuous auditing methodologies and technologies by researchers has been limited (Chan and Vasarhelyi 2011). For example, Alles et al. (2006) assisted in the implementation of an innovative monitoring and control layer for continuous monitoring of business process controls. Thiprungsri and Vasarhelyi (2011) introduced the use and application of cluster analysis to continuously monitor for irregularities in life insurance claims. Kim and Vasarhelyi (2013) applied unsupervised rules based learning methods to a wire payment transfer process to detect irregularities on a continual basis.

Thiprungsri and Vasarhelyi (2011) and Kim and Vasarhelyi (2013) show that a large volume of exceptions can be generated for investigation in the continuous auditing environment. In the research environment, the number of exceptions generated by a CA system may be an afterthought (Kuenkaikaew and Vasarhelyi 2013). However, in practice, a large volume of exceptions can incapacitate an internal audit department (Alles et al. 2006; Alles et al. 2008; Perols and Murthy 2012). Therefore, the management of exceptions is a critical problem in the CA environment because the investigation of exceptions is laborious and time consuming in nature. Limited resources in an audit department constrain the number of exceptions that can be investigated (Chan and Vasarhelyi 2011). This paper proposes a framework that attempts to maximize an audit department’s limited resources by prioritizing exceptions generated by a CA system. The framework utilizes the Dempster-Shafer theory of belief functions to systematically prioritize exceptions based on their suspiciousness.

Dempster-Shafer Theory of Belief Functions

The Dempster-Shafer theory of belief functions was introduced by Dempster (1967) and Shafer (1976). Belief functions utilize a constructive approach to explain evidence with probabilities. More specifically, this theory interprets the degree of belief through subjective probabilities and then utilizes Dempster's rule to combine the subjective probabilities from various sources. The use of belief functions in the auditing domain traces back to Shafer and Srivastava (1990). Srivastava and Shafer (1992) argued that belief functions are useful to represent the auditors' intuitive understanding of audit risk. Srivastava and Shafer (1992) linked the theory of belief functions to the structure of audit risk and presented formulas for audit risk with simple assumptions. More recently, Srivastava, Gao, and Gillett (2009) developed an algorithm based on belief functions to consider categorical and 'uncertain' logical relationships among binary variables.

The concept of prioritizing exceptions is not a new proposition. Walgampaya, Kantardzic, and Yampolskiy (2010) applied the Dempster-Shafer theory of belief functions to develop a data fusion mechanism for real time click fraud detection and prevention. For prioritization, the fusion mechanism combined multiple evidences to compute a suspicion score for each click (Hall and Llinas 1997). In the auditing domain, Perols and Murthy (2012) suggested further processing of detected anomalies using information fusion (White 1987). Information infusion is used to support exception detection, aggregation, and analysis of exceptions generated by the CA system.

Dempster-Shafer theory is a suitable information fusion method. Dempster-Shafer theory of belief functions can be used to develop data fusion mechanisms through combining evidence from multiple sources (Lefevre, Colot, and Vannoorenberghe, 2002; Walgampaya, Kantardzic, and Yampolskiy, 2010; Lelandais, Gardin, Mouchard, Vera, and Ruan, 2012). This study advances the continuous auditing literature by proposing a framework that utilizes belief functions to further process and prioritize identified exceptions in the CA environment.

1.3. Exception Prioritization Framework

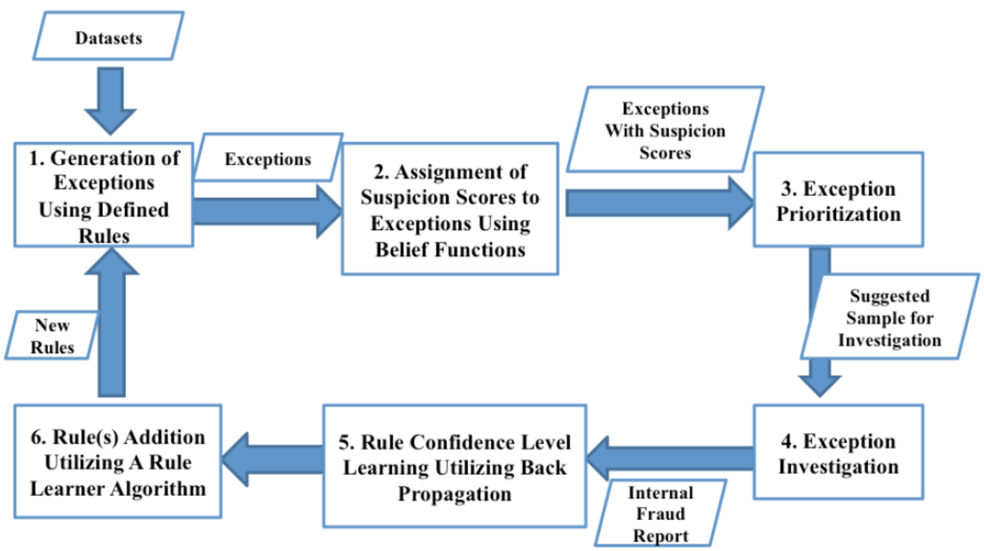

Figure 1.1: Exception Prioritization Framework

The main objective of this paper is to propose and validate a framework (Figure 1.1) that prioritizes exceptions in the CA environment. Prioritization of exceptions enables auditors to focus on those exceptions that are more likely to be irregular transactions. The framework has six stages: 1) generation of exceptions using defined rules, 2) assignment of suspicion scores to exceptions using belief functions, 3) exception prioritization, 4) exception investigation, 5) rule confidence level update utilizing back propagation, and 6) rule(s) addition utilizing a rule learner algorithm.

Stage 1: Generation of Exceptions Using Defined Rules

Expert knowledge is often translated into the form of if-then rules (Korver and Lucas 1993). Researchers in the accounting domain generally employ rule-based systems to detect irregularities (ex. errors, internal control violations, and/or fraud). For example, Deshmukh and Talluru (1998) applied a rule-based fuzzy reasoning system to evaluate the risk of management fraud. Lee (2008) integrated a rule-based approach with a case-based approach to detect internal control violations in a bank's internal auditing system. Kim and Vasarhelyi (2012) demonstrated an unsupervised rule-based model to detect fraud in an insurance company's wire payment transfer process.

In stage one of the proposed framework, expert rules are used to identify irregular transactions. The expert rules are initially defined by internal auditors. Each defined rule is assigned an initial confidence level based on the internal auditor's confidence that the rule can detect errors. In the framework, the accounting data is filtered through these rules. It is expected that a large number of exceptions will be generated by this process. Alles et al. (2008) and Perols and Murthy (2012) articulated that CA systems can inherently generate a large volume of exceptions. However, it is important to note that a transaction labeled as an exception may or may not indicate the existence of an error. These irregular transactions will only be labeled as errors if the auditor finds the instance to be during the investigation stage (stage 4).

Stage 2: Assignment of Suspicion Scores Using Belief Function

In stage two, the Dempster-Shafer theory of belief functions is used to assign suspicion scores for each transaction based on the defined rules that the transaction violated in stage one. This section demonstrates how the theory of belief functions is applied to represent uncertainties with expert rules in estimating the suspiciousness for each exception.

Belief functions have two main processes. The first is to generate the degree of belief for one hypothesis from the subjective probability and the second is to use Dempster's rule to combine the subjective probabilities from various sources. There are three important functions in Dempster-Shafer theory: the *basic probability assignment* (**m**), the *Belief* function (**Bel**), and the *Plausibility* function (**PL**). First, the basic probability assignment is used to present initial judgments. Second, the Belief function is used to express the suspicion scores of the exceptions. And third, the plausibility function is used to present the assignment of probability. Please refer to Appendix A for a simple example that demonstrates how the theory of belief function is used to estimate the suspicion score of a transaction.

The expert rules used in this study are assumed to be independent of each other to simplify the analysis. Each rule used for estimating the suspicion degree of an exception is treated as evidence. Both affirmative evidence and negative evidence are considered when the formulas are developed for estimating the suspicion degree of an exception. Affirmative evidence is defined as the evidence that supports a transaction to be irregular to a certain degree. Negative evidence is defined as the evidence that supports a transaction to be a normal transaction to a certain degree.

Dempster's rule is utilized to combine the confidence levels of the expert rules by assigning an aggregate suspicion score to an irregular transaction. The *Belief* function is interpreted as the suspicion score for transaction t, $\mathbf{Bel}_{\mathrm{t}}(f)$. The suspicion score is generated according to the work of Srivastava (2005). '$\sim f$' is interpreted to imply that the transaction is normal and 'f' is interpreted to imply that the transaction is an irregularity. Therefore, the entire frame is $\Theta = \{f, \sim f\}$.

Let us denote $m_i(f)$ the m-value of rule *i*. We have *n* different rules. If transaction *t* violates several rules, its suspicion score is given by:

$$\boldsymbol{Bel}_t(f) = m(f) = 1 - \prod_{i=1}^{n}\big(1 - m_i(f)\big)/K \qquad (1.1)$$

Since both the affirmative evidence and the negative evidence are considered, the renormalization K is given by

$$K = \prod_{i=1}^{n}\big(1 - m_i(f)\big) + \prod_{i=1}^{n}(1 - m_i(\sim f)) - \prod_{i=1}^{n} m_i(\{f, \sim f\}) \qquad (1.2)$$

Appendix A exhibits the formula derivation.

Each transaction will be assigned a suspicion score in this way. However, it does not mean that auditors will investigate all of the transactions with a non-zero suspicion score. Further processing to prioritize the exceptions based on various factors will be necessary.

Stage 3: Exception Prioritization

In stage three, the exceptions are ranked by their suspicion score and a threshold is defined to determine the investigative sample. The threshold is set by considering numerous factors such as the range of suspicion scores, the materiality of transactions, and the availability of audit resources (ex. time and labor). Only those exceptions with a suspicion score above the defined threshold will be taken into consideration for investigation. Figure 1.2 illustrates the exception prioritization schema.

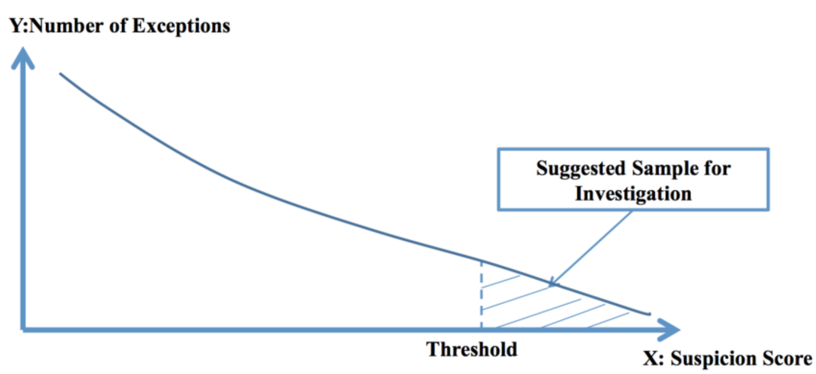

Figure 1.2: Prioritization Exceptions

In Figure 1.2, the x-axis represents the suspicion score and the y-axis represents the number of exceptions. There is an inverse relationship between the number of exceptions and suspicion score. Figure 1.2 illustrates that the number of exceptions decreases as the suspicion scores increases. This is consistent with the assumption that fewer transactions would violate multiple rules and thus there will be fewer transactions with higher suspicion scores. The suggested investigative sample is the shaded area under the curve.

Stage 4: Exception Investigation

The investigative element in auditing is human resource intensive and a time-consuming process (Hirst and Koonce 1996; Chan and Vasarhelyi 2011). The audit department will maximize audit resources by investigating those irregularities that have the highest likelihood of being errors. In stage four, the suggested investigative sample will be investigated manually by the auditors to determine if they are normal and erroneous transactions.

Stage 5: Rule Confidence Level Update Utilizing Back Propagation

Those irregular transactions that have been verified during investigation (stage four) as positive instances of errors are used to improve the effectiveness of the system in subsequent iterative cycles. The initial confidence levels assigned to individual rules are set arbitrarily based on the subjective judgment of internal auditors (stage one). In this stage, the confidence levels of these rules are adjusted using back propagation. Rules that contributed to finding errors are given a higher confidence level. In contrast, those rules that did not contribute to finding errors are given a reduced confidence level. When the confidence level of a rule decreases close to zero, the rule will be deleted from the set of rules. This process will increase the effectiveness of the framework by putting more weight on effective rules and less weight on ineffective rules in future iterative cycles.

The standard neural network learning algorithm called back propagation (Rumelhart and McClelland 1986) is utilized to refine the confidence levels of rules. The neural network learning algorithm is commonly used for revising belief in a network. For example, Towell and Shavlik (1994) proposed a knowledge-based artificial neural network that presents domain knowledge as a set of rules and refines this knowledge using back-propagation. Mahoney and Mooney (1993) utilized the standard formula for adjusting the weight between nodes when modeling the certainty-factor of summation evidence. According to Rumelhart and McClelland (1986), the standard formula for adjusting the weight (w_i) from the i th element of the input to the output after getting the real value, c, of the output is

$$\Delta w_i = \eta \delta O_i \tag{1.3}$$

where η is the adjustment rate pre-defined by the user, O_i is the value of the i th element of the input, and δ is the output error. The output error δ is defined as the difference between the calculated value of the output and its real value, $\delta = c - C$, where c is the real value for the output, and C is its assigned value.

The standard formula is used to model the revision of the confidence levels of the rules from stage one. In neural networks there are normally multiple levels that are used. However, in this study, all the rules are considered as elements at the same level. After investigation, if the internal auditor identifies transaction t as an error, the auditor revises the suspicion score for transaction t to be one or otherwise zero. Define $\Delta\ r_i$ as the adjustment of the confidence level for rule R_i on the basis of back propagation, and $\Delta_t r_i$ as the adjustment in terms of transaction t. Basically, the adjustment $\Delta\ r_i$ for rule R_i is the sum of the adjustments for all the transactions:

$$\Delta\ r_i = \sum_{t=1}^{n} \Delta_t r_i \tag{1.4}$$

The adjustments for the affirmative evidence and negative evidence are different. The formula derivation is presented in Appendix A. The adjustment of the confidence level r_i for affirmative rule R_i from the investigative finding of transaction t is:

$$\Delta_t r_i = \left(T_t - \mathbf{Bel}_{\mathrm{t}}(f)\right) * \frac{\prod_{j=1}^{n}\left(1-m_j(f)\right)m_i(f)}{K^2 * (1-m_i(f))}\left[\mathrm{K} + \prod_{j=1}^{n}\left(1-m_j(f)\right)\right] * \ r_i \qquad (1.5)$$

The adjustment of the confidence level r_i for the negative rule R_i from the investigative finding of transaction t is:

$$\Delta_t r_i = \left(T_t - \mathbf{Bel}_{\mathrm{t}}(f)\right)\left[-\frac{\prod_{j=1}^{n}\left(1-m_j(\sim f)\right) * \prod_{j=1}^{n}\left(1-m_j(f)\right)m_i(\sim f)}{K^2\left(1-m_i(\sim f)\right)}\right] * \ r_i \qquad (1.6)$$

where $\mathbf{Bel}_{\mathrm{t}}(f)$ is the suspicion score of transaction t that is assigned using belief functions in stage two and T_t is the real suspicion score for transaction t based on the investigative findings. If transaction t is error, T_t will take the value of 1, otherwise the value of 0. K is the renormalization that was defined in stage two. n includes these rules that the transaction violates.

The new confidence level of rule R_i will incorporate the prior confidence level and the aggregated adjustments from all the investigative transactions. Thus, the updated confidence level of rule R_i is defined as:

$$r_i{}' = r_i + \eta \Delta\, r_i \qquad (1.7)$$

where η is the pre-defined adjustment rate. The adjustment rate determines the degree of the rule's confidence level update.

Stage 6: Rule(s) Addition Utilizing a Rule Learner Algorithm

After each iterative run, a rule learner algorithm is used to develop additional rules. The rule learner algorithm generates new rules based on the positive instances of errors identified during the investigative stage (stage four). Although the confidence levels of the expert-based rules have been refined in step five, the newly identified erroneous transactions might have new attributes, which are not represented in the existing expert rules. The rule learner algorithm is used to convert these attributes into rule(s). The initial confidence level of the newly created rule(s) will be determined based on their ability to detect erroneous transactions.

The standard rule learner algorithm RIPPER is used as the rule learner in this study. Cohen (1995) developed RIPPER and it is an extension of the IREP rule-learning algorithm introduced by Furnkranz and Widmer (1994). The RIPPER algorithm generates rules in two stages: the growing stage and the pruning stage. The current set of training instances is divided into two subsets, a growing set and a pruning set. In the growing stage, rules are constructed from the instances in the growing set. The new rules are developed through the exhaustion of all the possible values and combinations of attributes in relations to the labeled value (ex. regular or erroneous transaction). In the pruning stage, some rules will be removed in order to improve the performance of the rules on the pruning instances. Rules are pruned and replaced with the metric of minimizing the errors in the entire rule set on the pruning set.

The RIPPER algorithm has been wildly used as the standard rule-learner method in the machine learning literature. One advantage of the RIPPER algorithm is its ability to build compact and understandable rules that can be generally interpreted by humans and thus can be interpreted by auditors. Furthermore, the RIPPER algorithm has been proven to work well on large and noisy datasets (Cohen, 1995; Sasaki and Kita, 1998). However, the standard RIPPER algorithm is not cost-sensitive and does not support the assignment of misclassification cost. Moreover, RIPPER does not include a mechanism to address the problem with data that has a skewed class distribution. Pietraszek (2004) refined RIPPER using weighting to fulfill the cost-sensitivity and incremental learning. Another issue with the RIPPER algorithm is its inability to keep a condition that appears in more than one category. If there is such a condition, RIPPER will remove the rule in the pruning stage.

1.4. Data

Data Sample

Accounts payable transactional data from a technology company was used in the experiment. Due to the company's sensitivity in disclosing actual transactional errors, the study simulates irregularity labeling of these transactions using a novel approach. Erroneous transactions are simulated using k-means clustering. The anomalous transactions identified in clustering are assumed to be erroneous transactions. Appendix B provides details about the identification and labeling of irregular transactions using k-means clustering.

The period of the data ranges from 2000 to 2010 and consists of 89,712 transactions. Descriptive statistics for the dataset are presented in Table 1.1.

Table 1.1: Descriptive Statistics of Accounts Payable Data

Variable	Category	Mean	Minimum [a]	Maximum [a]
Vendor_ID	Character	N/A	N/A	N/A
Vendor_Name	Character	Missing	Missing	Missing
Invoice_No	Character	N/A	N/A	N/A
Voucher_Description	Character	N/A	N/A	N/A
Invoice_Date	Numeric	3/31/09	1/13/00	6/24/10
Amount	Numeric	8,073.59	-46,656.50	3,435,664.00
Tax_Amount	Numeric	308.30	-4,241.50	44,000.00
Goods_Amount	Numeric	7,765.26	-42,415.00	3,435,664.00
Voucher_No	Character	N/A	N/A	N/A
Invoice_Type	Character	N/A	N/A	N/A
Due_Date	Numeric	3/15/10	2/12/00	8/17/10
Full_Pay_Status	Character	N/A	N/A	N/A
Date_Full_Payment_Due	Numeric	4/9/10	4/1/09	6/28/10
Payment_Date	Numeric	4/30/10	10/31/08	6/30/10
Batch Number	Numeric	N/A	N/A	N/A
GL Account	Character	N/A	N/A	N/A
Bank ID	Character	N/A	N/A	N/A
Payment ID	Numeric	N/A	N/A	N/A

[a]If the variable is in the date format, then the minimum is the first date and the maximum is the last date.

The dataset is randomly partitioned into a training subset and a testing subset (Table 2). There are 53,828 (60%) transactional observations in the training subset and there are 35,884 (40%) transactional observations in the testing subset (Table 1.2). In regards to the individual training and testing subsets, there are 947 erroneous transactions in the training subset and 618 erroneous transactions in the testing subset. The total number of regular transactions in both subsets is 88,147 (98.3 %). There are 52,881 regular transactions in the training subset and 35,266 regular transactions in the testing subset.

Table 1.2: Data Sample Assignment

	Training Subset	Testing Subset	Total
Regular Transactions	52,881	35,266	88,147 (98.3%)
Labeled Errors	947	618	1,565 (1.7%)
Total	53,828 (60%)	35,884 (40%)	89,712

Expert Rules

The internal auditors at the technology company developed rules for detecting irregular transactions in accounts payable data. Table 1.3 presents the thirty rules developed by the technology company's internal auditors. Each rule was assigned a weight or confidence level of low, medium, or high by the auditors to represent its relative importance or confidence in identifying errors. In the experiment, the confidence levels of low, medium, and high were converted to the numerical values of 0.25, 0.5, and 0.75, respectively. These numerical values are treated as the initial confidence level of the rules. There were three rules that were not assigned a confidence level by the auditors. These rules were arbitrarily assigned a confidence level of low (0.25).

Table 1.3: Expert-based Rules Overview

Rules	**Initial Confidence Level**
Disbursements Posted without Invoices	High
Invoices with Vendors that Do Not Appear on the Vendor Master List	High
Payment is a Negative Amount	High
Payment is a Zero Dollar Amount	High
Keywords Search	Medium
Outlier Analysis - Disbursements to Vendor	Medium
Outlier Analysis - Disbursements by G/L Account	Medium
Payments Made on the Weekend	Medium
Payments Made on a Holiday	Medium
Missing Disbursement Date	Low
Missing Invoice Type	Low
Invoices with invalid GL account information	Low
Round Amount Disbursements (by line item)	Low
Round Amount Disbursements (by invoice)	Low
Duplicate Invoices/Disbursements	Low
Vendors with Multiple Invoices Per Day	Low
Payment Date vs. Due Date Analysis	N/A [a]
Payment Date vs. Invoice Date Analysis	N/A [a]
Gaps in Voucher Number Sequence	N/A [a]

[a] N/A: Internal auditors did not provide a confidence level for the rule.

1.5. Experiment Design

The framework consists of two main processes, Process A and Process B (Figure 1.3). Process A involves the first three stages of the framework; 1) generation of exceptions using defined rules, 2) assignment of suspicion scores to exceptions using belief functions, 3) exception prioritization. Process A is used to generate and prioritize exceptions for investigation. Process B involves the last three stages of the proposed exception prioritization framework; 4) exception investigation 5) rule confidence level update utilizing back propagation and 6) rule(s) addition utilizing a rule learner algorithm.

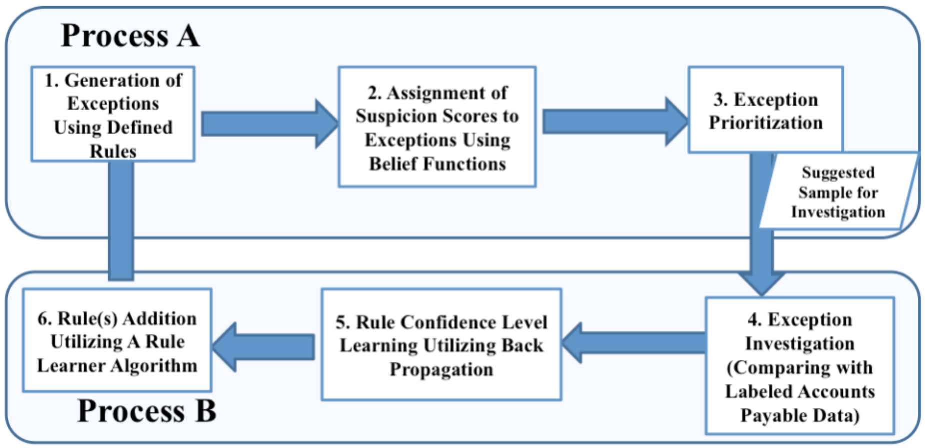

Figure 1.3: Experiment Process Diagram

Process A: Investigative Sample Generation

In Stage 1, the accounts payable transaction data is filtered through a set of rules created by internal auditors (expert rules) to identify exceptions. The internal auditors using professional judgment assign a confidence level to each rule to indicate its ability to detect errors. In Stage 2, each transaction is assigned a suspicion score through applying the belief function schema. In Stage 3, transactions are ordered or prioritized based on their assigned suspicion score. Then, the top ten percent of ranked prioritized exceptions are chosen for investigation in the experiment. In subsequent iterative run, the size of the investigative sample is arbitrarily fixed at ten percent to evaluate the performance of the framework in prioritizing errors. This results in 5,382 transactional observations in our investigative sample for the training process, and 3,588 transactional observations in the investigative sample for the testing process. The investigative sample size is determined without consideration of materiality to simplify the experiment.

Process B: Comparison and Learning

Process B in the experiment compares the suggested investigative sample from Process A with the labeled erroneous transactions identified from clustering. After comparing the investigative sample with the labeled account payable data, Process B involves the last two stages of the framework; Stage 5) rule confidence level updating utilizing back propagation and Stage 6) rule(s) addition utilizing a rule learner algorithm.

Back propagation is used to update the confidence levels of the rules. The experiment follows Formula 1.5 and Formula 1.6 (see above Stage five) of the framework to refine the confidence levels of the rules. The adjustment rate η is defined as 25 in this experiment.

When using any type of machine learning algorithm such as the rule learner RIPPER, researchers must consider the methodological concern of an unbalanced dataset. Normally, there are many more instances of regular transactions than there are errors. An imbalanced dataset can make an algorithm ineffective when trying to capture the characteristics of the erroneous transactions. The combination method was used to deal with the imbalance issue in the study. The choice of the method will depend on the performance of the different methods during testing. Therefore, the method chosen in this study may not be a universally preferred approach in all situations. Appendix B provides a detailed explanation on how this study chose the combination method to solve the imbalance issue.

The objective of the rule learner algorithm is to discover hidden rules that can be used to predict an outcome (Error). The rule learner algorithm identifies rule(s) that capture the characteristics of erroneous transactions using the training dataset. In the experiment, the new rules are created using all the observations in the training subset to simplify the experiment. The new rule(s) are evaluated using the testing dataset and those rules that are effective in identifying erroneous transactions are added to the original expert rules for use in future iterative runs. The new rule(s) are assigned a confidence level of high, medium, or low depending on the new rule(s) ability to detect errors in the testing dataset. The conversion of the confidence levels to numeric values is similar to how it was applied in converting the confidence levels of the original expert rules. Furthermore, the confidence levels of these new rules will be revised using back propagation in subsequent iterative runs. The rule learner algorithm RIPPER was used in this study and it generated 30 new rules. Table 1.4 lists these new rules and their initial confidence levels.

Table 1.4: New Rules Generated Using the Rule Learner Algorithm

Rule Number	Rule Details	Initial Confidence Level
1	(Amount <= 1480) and (Amount >= 1396.4) and (TotalAmount <= 25756.2) and (GAP_Voucher_No = 1)	High
2	(Amount <= 776.49) and (Amount >= 677.05) and (NO_INVOICE >= 17) and (Amount >= 708.37) and (Amount <= 732.91)	High
3	(Invoice_No = AUS-10-000832)	High
4	(Totalcredit <= -6.6) and (Amount <= 3117.62) and (NO_INVOICE >= 26) and (Payment_Due_Date = 0) and (Invoice_No = AUS-10-002806)	High
5	(Totalcredit <= -6.6) and (Amount <= 4331.56) and (GAP_Voucher_No = 0) and (Invoice_No = AUS-10-001733)	High
6	(TotalAmount >= 70621.35) and (TotalAmount <= 88940.77) and (Invoice_No = AUS-11-000490)	High
7	(NO_INVOICE >= 30.418833) and (TotalAmount <= 16618.563965) and (Amount >= 556.988653)	High
8	(NO_INVOICE <= 4) and (Amount <= 792) and (Amount >= 690.8)	High
9	(Invoice_No = AUS-11-000466)	High
10	(Amount <= 4634.4) and (TotalAmount >= 62508.36) and (Totalcredit >= -25.381926) and (NO_INVOICE <= 20)	Medium
11	(MultiInvoice_Per_Day = 0) and (Amount >= 2584.347345) and (Amount <= 4059.85)	Medium
12	(Amount <= 1678.08) and (Amount >= 1376.1) and (Amount <= 1460.61) and (TotalAmount <= 25116.94)	Medium
13	(Invoice_No = AUS-10-000355)	Medium
14	(TotalAmount >= 71482.69) and (Amount <= 2867.33) and (NO_INVOICE <= 46.91419) and (NO_INVOICE >= 35) and (Payment_Due_Date = 0)	Medium

15	(Amount <= 905.272033) and (NO_INVOICE >= 30.418833) and (NO_INVOICE <= 31) and (Amount >= 672.1)	Medium
16	(MultiInvoice_Per_Day = 0) and (Amount >= 14440.8) and (Amount <= 15529.44)	Medium
17	(NO_INVOICE <= 9.581865) and (TotalAmount >= 19907.85) and (TotalAmount <= 22905.4)	Medium
18	(Invoice_No = AUS-10-002039)	Medium
19	(Invoice_No = AUS-10-001522)	Medium
20	(Invoice_No = AUS-11-000289)	Medium
21	(NO_INVOICE >= 48) and (Payment_Due_Date = 0)	Low
22	(Amount <= 805.2) and (TotalAmount >= 45533.75) and (NO_INVOICE <= 42)	Low
23	(MultiInvoice_Per_Day = 0) and (TotalAmount <= 3738.6) and (Outlier_Disbur = 1)	Low
24	(TotalAmount >= 73270.36) and (Totalcredit >= -25.381926) and (TotalAmount <= 77393.922903)	Low
25	(TotalAmount >= 41312.459993) and (Invoice_No = 901808675/SEPT09)	Low
26	(Vendor_ID = BAK001)	Low
27	(Vendor_ID = P-THE002)	Low
28	(Invoice_No = AUS-11-000268)	Low
29	(Vendor_ID = X-HIGCH01)	Low
30	(Vendor_ID = P-CHI003)	Low

1.6. Experiment Results

This section discusses the experiment conducted to evaluate the performance of the exception prioritization framework

In the training process, the framework is trained using twenty cycles. The performance of the framework improved after each iterative run until reaching a plateau and it approximately converged in the sixth iterative cycle. The framework is then subsequently evaluated on the testing dataset. Table 5 summarizes the suspicion scores per cycle for both the training subset and the testing subset. The testing results indicate that there is no over-fitting issue. The amount of errors that was correctly identified in the testing process does not dramatically decrease with each iterative run. Furthermore, the results from the testing process are very similar to the training process. Table 5 shows the training and testing process reached a plateau after the fourth iterative cycle. After the thirteenth iterative cycle, the performance of the framework depleted which indicates over-training (Table 5).

Table 1.5: Framework Evaluations

Cycle	Number of Detected Errors [a]		Percentage of Detected Errors [b]		Mean of Suspicion Score of Errors [c]		Mean of Suspicion Score of Normal/Regular Transactions [d]	
	Train	Test	Train	Test	Train	Test	Train	Test
1	473	265	49.95%	42.88%	0.732	0.71	0.568	0.569
2	557	344	58.82%	55.66%	0.625	0.61	0.449	0.449
3	577	352	60.93%	56.96%	0.606	0.593	0.449	0.449
4	594	369	62.72%	59.71%	0.597	0.584	0.449	0.449
5	594	369	62.72%	59.71%	0.591	0.579	0.449	0.449
6	595	368	62.83%	59.55%	0.587	0.576	0.45	0.45
7	595	368	62.83%	59.55%	0.584	0.573	0.45	0.45
8	595	368	62.83%	59.55%	0.582	0.571	0.451	0.45
9	595	368	62.83%	59.55%	0.581	0.57	0.451	0.451
10	595	368	62.83%	59.55%	0.579	0.569	0.451	0.451
11	595	368	62.83%	59.55%	0.578	0.568	0.451	0.451
12	595	368	62.83%	59.55%	0.577	0.567	0.451	0.451
13	595	368	62.83%	59.55%	0.576	0.566	0.452	0.451
14	577	351	60.93%	56.80%	0.575	0.566	0.452	0.452
15	594	370	62.72%	59.87%	0.574	0.565	0.452	0.452
16	576	353	60.82%	57.12%	0.573	0.564	0.452	0.452
17	576	353	60.82%	57.12%	0.573	0.564	0.452	0.452
18	576	353	60.82%	57.12%	0.572	0.563	0.452	0.452
19	576	353	60.82%	57.12%	0.571	0.563	0.452	0.452
20	576	353	60.82%	57.12%	0.571	0.562	0.453	0.452

[a] The number of errors in the investigative sample for the training subset and the testing subset.

[b] The percentage of errors that are correctly identified for the training subset and the testing subset.

[c] The mean of the suspicion score for all the errors for the training subset and the testing subset.

[d] The mean of the suspicion score for all the normal transactions for the training subset and the testing subset.

The experiment relies on two criteria to evaluate the performance of the framework. The first one is the ability to effectively prioritize erroneous exceptions higher than non-erroneous exceptions. The second criterion is the framework's ability to improve its prioritization performance after each iterative run. In terms of the first criterion, the results indicate that on average the erroneous transactions were assigned a higher suspicion score than regular transactions after each iterative run (Table 1.5). This indicates that the proposed framework has the ability to prioritize the erroneous exceptions higher than other exceptions. It is observed that the mean of the suspicion scores for the erroneous transactions is higher than the mean for the regular transactions in each iterative cycle. In terms of the second criterion, the suspicion scores are compared across cycles. It is also observed that the differences of the mean of the suspicion scores between each cycle for the erroneous transactions are always less than the regular transactions. Collectively, these findings show that erroneous transactions are being prioritized higher than regular transactions and this test validates the effectiveness of the framework.

Using rule confidence level updating, the results show that the performance of the framework improved in prioritizing exceptions after each iterative run. In the first run, 42.88 percent of erroneous transactions are prioritized into the investigative sample. This percentage increases after each iterative run until reaching a peak of 59.71 percent in the fourth iterative run. The increased percentage in prioritizing erroneous transactions higher than regular transactions for the investigative sample indicates the effectiveness of back propagation in the framework.

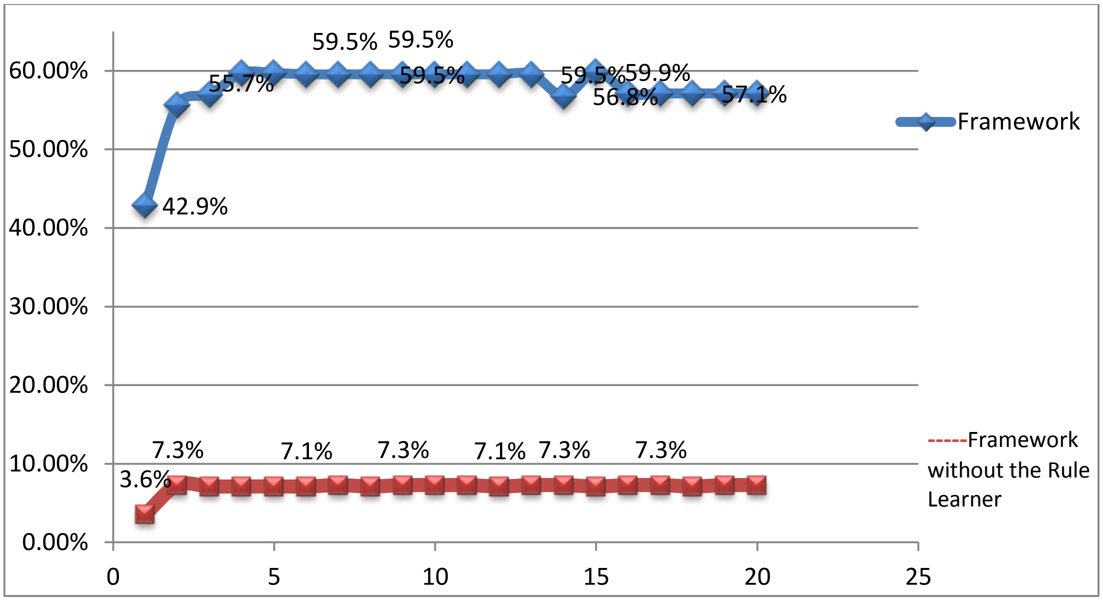

Figure 1.4: Compare the Framework to the Framework without the Rule Learner Algorithm

Note: The y-axis presents the percentage of labeled errors that were correctly identified. The blue line shows the performance of the system with the rule learner algorithm, and the red line shows the performance of the system without the rule learner.

Figure 1.4 presents the performance of the system applying the rule updating algorithm and the system not applying the rule updating algorithm. The former consists of the expert rules and the new rules that were generated using the rule learning algorithm. The latter only consists of the expert rules. Although both of the systems improved through the use of rule confidence level updating, the system with the new rules outperforms the one without that feature. In the first run, the system with the additional new rules detected 42.88 percent of the erroneous transactions. Comparatively, the one without the new rules detected 3.56 percent of the erroneous transactions. In the fifth run, the one with the new rules detected 59.87 percent of the erroneous transactions and the one without the new rules identified 7.12 percent of the erroneous transactions. The performance of the system without the new rules is limited by the expert rules. This is because the experiment uses artificial errors, and the original expert rules created by the internal auditors were not designed to detect them. The experiment results show that the rule learner algorithm feature improved the framework's effectiveness by adapting to the artificial errors at hand.

The framework utilized Dempster-Shafer theory of belief functions to generate a suspicion score for each transaction in stage two. It can be debated that probability theory can be utilized to generate suspicion scores as well. Probability theory is used as a benchmark to see whether its performance outperforms that of Dempster-Shafer theory of belief functions. Figure 1.5 presents the results of the framework using belief functions and the results of the framework using probabilities. The results show that belief functions can be a better way to generate suspicion scores.

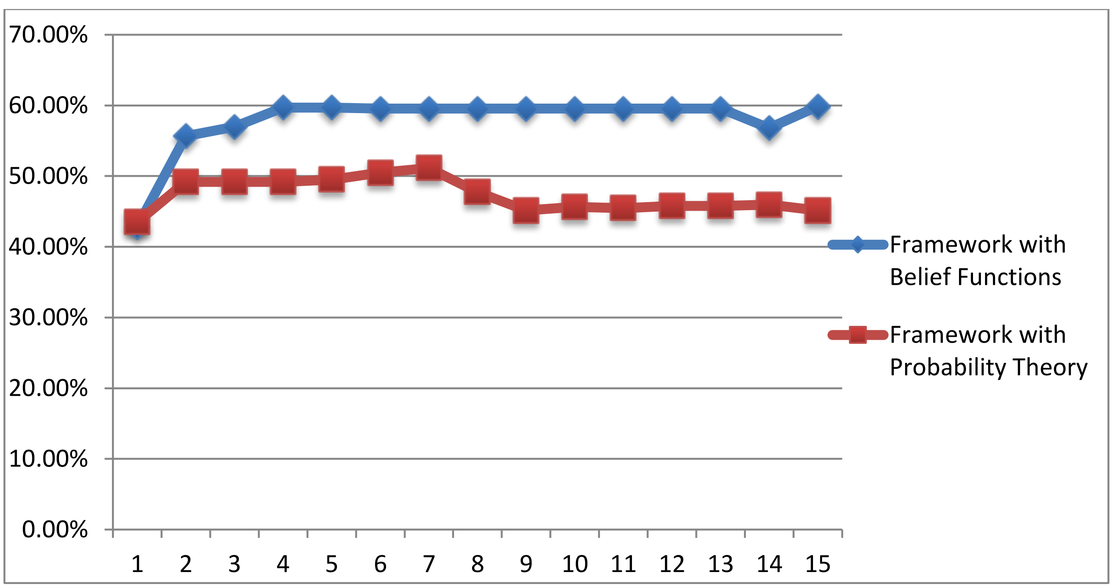

Figure 1.5: Framework with Belief Functions vs Framework with Probability

Note: The y-axis presents the percentage of labeled errors that were correctly identified. The blue line shows the performance of the framework with belief functions, and the red line shows the performance of the framework with probability theory.

1.7. Robustness

For robustness and sensitivity testing, the adjustment rate, initial confidence levels, and the parameters in clustering were set to different values.

Adjustment Rate

The adjustment rate η in stage five determines the degree of change for the confidence level of the rules. Utilizing a sensitivity test, the adjustment rate was set to the following values; 20, 30, and 35. Table 1.6 presents the percentage of the errors that was correctly identified in the test process when the system uses these adjustment rates. The results indicate that the framework performs similarly with these different adjustment rates and does not contrast the original results.

Table 1.6: Back Propagation Applying Different Adjustment Rates

Cycle	Percentage of Detected Errors			
	Adjustment Rate 20	Adjustment Rate 25	Adjustment Rate 30	Adjustment Rate 35
1	42.88%	42.88%	55.66%	42.88%
2	55.99%	55.66%	42.88%	58.74%
3	59.39%	56.96%	55.66%	53.07%
4	60.19%	59.71%	56.80%	53.07%
5	60.03%	59.71%	57.12%	53.07%
6	60.03%	59.55%	56.80%	53.07%
7	60.03%	59.55%	56.80%	53.07%
8	60.03%	59.55%	59.55%	51.13%
9	60.03%	59.55%	59.55%	53.07%
10	60.03%	59.55%	56.80%	52.10%

Confidence Levels

It can also be debated that the initial numeric confidence levels may influence the performance of the framework. To address this concern, the initial confidence levels of the expert rules were adjusted to test sensitivity. The low, medium, and high confidence level values were set to the numerical values of 0.2, 0.4, and 0.7, respectively. For the rules without a confidence level assigned by the auditors, their initial confidence level was set to 0.3. Table 1.7 presents the results when the system was trained with these initial confidence levels. The results still indicate that framework is effective in prioritizing erroneous transactions.

Table 1.7: Systems with Different Initial Confidence Levels

Cycle	Number of Detected Errors [a]		Percentage of Detected Errors [b]	
	Train	Test	Train	Test
1	466	270	49.21%	43.69%
2	550	336	58.08%	54.37%
3	568	344	59.98%	55.66%
4	568	347	59.98%	56.15%
5	568	347	59.98%	56.15%
6	567	346	59.87%	55.99%
7	567	346	59.87%	55.99%
8	570	344	60.19%	55.66%
9	569	344	60.08%	55.66%
10	569	344	60.08%	55.66%
11	568	345	59.98%	55.83%
12	563	343	59.45%	55.50%
13	563	343	59.45%	55.50%
14	564	345	59.56%	55.83%
15	563	343	59.45%	55.50%

[a] The number of errors that was prioritized into the investigative sample for the training subset and the testing subset, respectively.
[b] The percentage of the number of errors that are correctly identified for the training subset and the testing subset, respectively.

Clustering Parameters

Artificial errors were generated using a clustering algorithm in this experiment. There may be several concerns about this approach in generating simulated errors. First, the number of clusters is arbitrarily chosen. To test whether a change in the number of clusters will influence the results, the second optimal number of clusters (seventeen) is used. Second, the probability of the observation belonging to a cluster is arbitrarily chosen. A transaction is treated as an anomaly if the probability of that transaction is less than the arbitrarily defined threshold. To test whether a change in the probabilities will influence the results, those transactions with probabilities of less than 0.55 were treated as anomalies. Third, the data sample has been separated into 75:25 ratio used for training and testing. Figure 1.6 shows this change in parameters does not change the performance of the framework and provides consistent results.

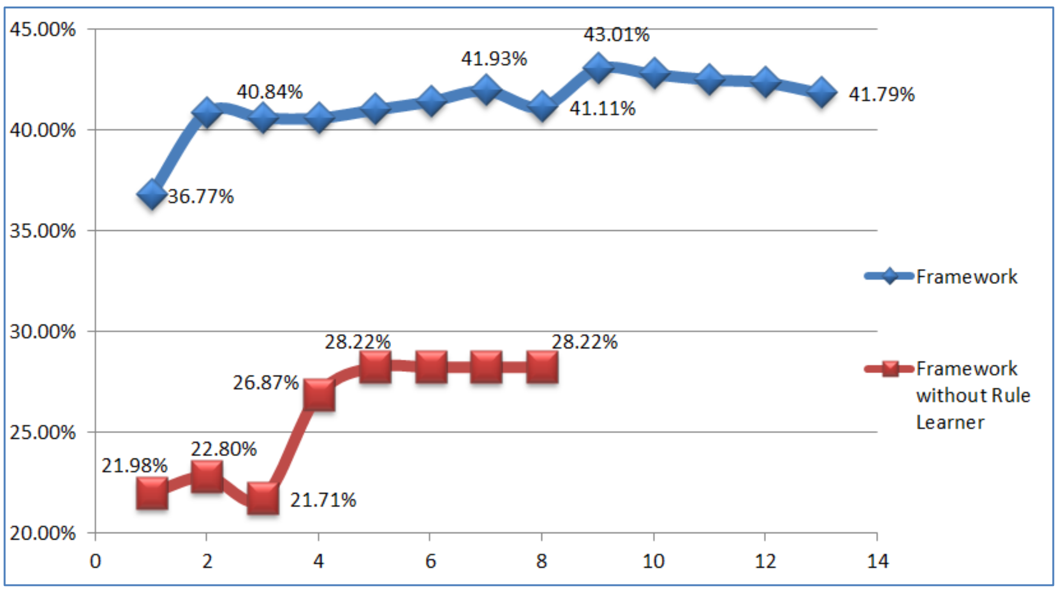

Figure 1.6: Testing Results with New Artificial Errors

Note: The y-axis presents the percentage of labeled errors that were correctly identified. The blue line shows the performance of the system with the rule learner algorithm, and the red line shows the performance of the system without the rule learner.

1.8. Contribution and Future Work

Researchers and practitioners have found that the volume of exceptions generated in a CA system can overwhelm an audit staff and may diminish any efficiency gains due to automation. To utilize audit resources more effectively, a framework is proposed to systematically prioritize exceptions generated by a CA system. The proposed prioritization framework is a semi-automatic system. The framework only relies on human experts to 1) initially define rules and assign confidence levels to the rules and 2) to subsequently investigate the prioritized exceptions. The rest of the framework is automated by the CA system. A prominent feature of the framework is the use of back propagation to increase the system's accuracy in prioritizing erroneous transactions higher than normal transactions after each iterative process. Furthermore, a rule learner algorithm feature is implemented to generate new rules that identify additional characteristics of errors.

The proposed framework is evaluated using a simulated experiment. The results from the experiment provide supporting evidence that the framework is effective in prioritizing exceptions and can learn from each iterative run. However, the experiment performed in the study has a number of limitations. First, erroneous transactions were artificially simulated. These erroneous transactions may not accurately reflect real world examples of errors. Second, each rule is assumed to work in isolation from the other rules. The effect of multiple rules adds complexity to the system and was not implemented in the experiment.

Third, the information from the investigative results is limited since it includes only a portion of all accounts payable transactions. As a result, the framework misses information about irregularities that are not present in the sample. Chychyla and Kogan (2014) suggested using a statistical model to capture the underlying distribution of the transactional data to mitigate this feedback problem. Forth, an implicit assumption is made that positive erroneous transactions will be identified when flagged by the system so that the belief function will be updated appropriately. However, we understand that auditing will not always return the "correct" results and can miss erroneous transactions, thus negatively affecting system updates. Fifth, the framework was applied to the accounts payable cycle. The framework may not be as effective in other business cycles.

For future research, the interrelationship between rules can be considered. The interaction of rules may help improve the effectiveness of the framework in prioritizing exceptions. In practice, the rules may have complex inter-relationships and may affect the outcome of the network. More advanced belief networks can be used to represent the dependencies and independencies among these rules (Korver and Lucas 1993). Researchers may also want to apply or extend the framework to other process audits such as revenue, inventory, payroll, and other cycles to verify the universal applicability of this framework.

Chapter 2 Application of Customer Search Volume In Analytical Procedures

2.1. Introduction

Audit standards (AICPA 1988, 2002) encourage external auditors to consider both financial information and nonfinancial information when performing analytical procedures and assessing fraud risk. Analytical procedures using only financial information are believed to be ineffective for evaluating the validity of financial statement data and detecting fraud (PCAOB 2004, 2007). Nonfinancial information is considered to be relatively free from the manipulation of managers and useful for analytical procedures (Brazel, Jones, and Zimbelman 2009; Ames et al. 2012). Research indicates that auditors have recently gathered and considered more nonfinancial information that is available through the Internet (Trompeter and Wright 2010). Previous literature has examined various types of nonfinancial information that can improve the prediction of analytical procedures, such as the economic and industrial data (Lev 1980; Loebbecke and Steinbart 1987; Wild 1987) and customer satisfaction (Ittner and Larcker 1998). Brazel, Jones, and Zimbelman (2009) indicate that nonfinancial information such as corporate operational measures can improve fraud detection.

This study investigates whether the consumer search volume can be used as nonfinancial information in analytical procedures to improve both the effectiveness of the expectation and error or fraud detection. The consumer search volume is the aggregate number of times that consumers submit terms to search engines. The extant research argues that the consumer search volume can capture the general level of consumer interest in terms of corporate products or services (Kulkarni, Kannan, and Moe 2012; Du and

Kamakura 2012). This proxy for consumer interest can be correlated with retail sales. The consumer search volume can allow auditors to analyze the trend of consumer interest and improve analytical procedures. During analytical procedures, auditors compare the client's financials with the expectations generated by their own models. The expectation model incorporating the consumer search volume is expected to generate more accurate predictions than the models that do not utilize this information. The expectation model incorporating the consumer search volume is also expected to improve fraud or error detection.

This study uses an empirical approach to examine whether the consumer search volume improves prediction and error detection in analytical procedures. The data for the consumer search volume are obtained from Google. Google is the largest search engine by market share[1]. The search volume gathered by Google likely represents the search behavior of the general population on the Internet. Recent studies have used the search volume index (SVI) reported by Google. Ginsberg et al. (2009) indicates that the SVI predicts flu outbreaks 1 to 2 weeks earlier than the US Center for Disease Control and Prevention (CDC). The extant literature in economics has examined the application of the SVI in the analysis and forecast of economic indicators and found that the models with the SVI can provide more accurate predictions than conventional models (Askitas and Zimmerman 2009; Vosen and Schmidt 2011; Wu and Brynjolfsson 2013; Li et al. 2015; Wu and Deng 2015). Research in finance has indicated that the SVI can be a direct proxy of investor sentiment (Joseph, Wintoki, and Zhang 2011; Da, Engelberg, and Gao 2015)

[1] Google generated 68.8 percent of all core search queries in the United States and accounted for 89.3 percent of the global search market as of mid-2015.
http://www.statista.com/statistics/267161/market-share-of-search-engines-in-the-united-states/

and of the information demand of individual investors (Vlastakis and Markellos 2012; Drake, Roulstone and Thornock 2012). Da, Engelberg, and Gao (2011) show that the SVI captures investor attention in a timelier manner than existing proxies such as extreme returns, trading volume, news and headlines. Following these studies, various other papers in the finance and accounting literature have also used the SVI as a proxy for investor attention (Bank, Larch, and Peter 2011; Andrei and Hasler 2014; deHaan, Shevlin, and Thornock 2015; Goddard, Kita, and Wang 2015). Overall, the extant research confirms the validity of the SVI. Therefore, this study uses the SVI reported by Google.

Google provides Search Volume Index (SVI) for search terms through its product, Google Trends. Google Trends returns the monthly or weekly SVI for search terms. A search term may have multiple meanings (e.g., "Apple"). To address this issue, Google Trends is featured with categories to refine the search results. For example, the search of "McDonald's" can be limited to the "Food and Beverage" main category, which can be narrowed down to the "Restaurant" subcategory, or even further refined to the "Fast Food" sub-subcategory. Figure 2.1 plots the weekly SVI of two terms from January 2004 to December 2015. The first graph displays the SVI of "Blackberry" in the "Mobile & Wireless" category and the total revenue of Blackberry Limited. As displayed in Figure 2.1, the SVI of "Blackberry" spikes around 2011 and falls afterwards. This trend is consistent with the pattern of the total sales of Blackberry Limited. The second graph displays the SVI of "Nike" in the "Athletic Apparel" category and the total revenue of Nike, Inc. The SVI of "Nike" increases during the holiday season (November and December), which is consistent with the notion that consumers are more interested in buying Nike products during the Thanksgiving and Christmas holidays.

The pattern of this SVI is consistent with the trend of the total revenue of Nike, Inc.

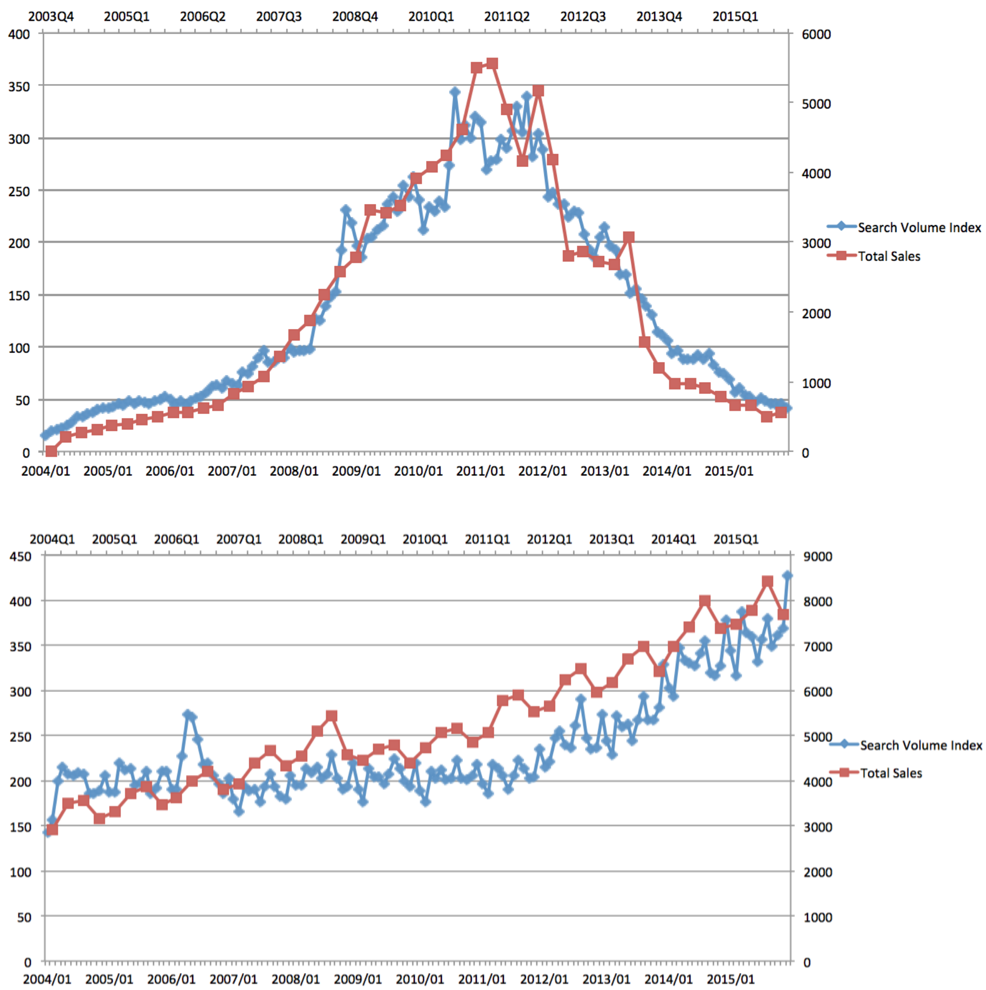

Figure 2.1: Search Volume Index and Total Sales

NOTE: The first graph displays the search volume index of “Blackberry” and the total sales of Blackberry Limited. The second graph displays the search volume index of “Nike” and the total sales of Nike. Inc

This study focuses on a sample of firms from 23 retail and manufacturing industries. The sample is restricted to firms that manufacture or sell products to consumers. This study builds a custom list of search terms for each firm. To capture the consumer search behavior related to each firm, the study uses corporate brands as search terms. When using corporate brands as search terms, there are some generic terms even though the search results are refined using categories in Google Trends. To address this issue, any search term that is generic under its specific category is excluded from the sample. The SVI data are collected for all the firms under the sample. Then, the SVI data are merged with corporate financial data. Previous studies show that the disaggregated monthly financial data outperforms quarterly financial data in analytical procedures (Wild 1987; Dzeng 1994; Chen and Leitch 1998, 1999; Leitch and Chen 2003; Hoitash, Kogan, and Vasarhelyi 2006). Thus, this study employs monthly financial data, which is interpolated from quarterly financial data.

This study uses the consumer search volume measured by the SVI to generate the predictions for the total revenue in analytical procedures. The generated predictions are evaluated using the criteria of mean absolute percentage error (MAPE), which is calculated using the absolute difference between the predicted value and the actual value. A smaller value of MAPE indicates better performance. Based on this criteria, the results indicate that the predictions that are generated using the consumer search volume statistically outperforms the predictions that do not use the consumer search volume for 17 industries out of 23 selected industries in the sample.

This study further examines whether the consumer search volume can improve the detection of fraud or errors. The inconsistencies between financial and nonfinancial measures can be pronounced for firms with fraud or errors. For example, increased retail sales accompanied with decreased consumer search volume may raise a potential red flag for auditors. To evaluate the performance of the consumer search volume on error detection, this study uses a simulated experiment. First, this study examines the error detection ability of the model that incorporates the consumer search volume when uncoordinated errors occur. In this scenario, simulated errors only seed into the dependent variable (e.g. overstated total revenue). The model that incorporates the consumer search volume is compared with the benchmark model with respect to error detection. The results indicate that the model with the consumer search volume results in less false positives and false negatives than the benchmark model for 17 out of 23 selected industries. Second, this study examines the error detection ability of the model that incorporates the consumer search volume when coordinated errors occur. In this scenario, simulated errors seed into the dependent variable and the independent variable (e.g. overstated total revenue and overstated accounts receivable). The results show that the model with the consumer search volume generates less false positives and false negatives than the benchmark model for 13 out of 23 selected industries.

This study makes the following contributions to the literature. First, researchers have used the SVI in various areas, including economics, finance, marketing, and accounting. This is the first study that applies the search volume into the auditing area. Second, this study considers the search volume from the perspective of consumers and applies it to analytical procedures. The quantitative feature of the search volume allows auditors to analyze the pattern and change of consumer interest.

The rest of the paper is organized as follows. Section 2.2 provides a review of the relevant literature and research questions. Section 2.3 describes the data and methodology. Results are shown in Section 2.4. Section 2.5 concludes and provides the limitation of this study.

2.2. Background and Research Questions

The section reviews these studies that have been done to improve analytical procedures, and then reviews the studies that use the search volume index reported by Google. Lastly the section discusses three questions about examining the performance of the consumer search volume as nonfinancial information in analytical procedures.

Analytical Procedures

Analytical procedures include comparison of the clients' accounts balance with the expectation that generated by auditors through analysis of plausible relationships among both financial and nonfinancial data. SAS No. 56 provides detailed guidance on using analytical procedures as substantive tests and requires auditors to use analytical procedures both in the planning and overall stage of all audits (AICPA 1989). SAS No. 56 suggests that nonfinancial information should be considered when performing analytical procedures, and used to evaluate risks and detect material misstatements (AICPA 2002, 2007). The PCAOB believes that nonfinancial information is more powerful and reliable for the evaluation of financial statement (PCAOB 2004, 2007).

Extensive research has been done to improve the efficiency and effectiveness of analytical procedures in auditing. Research about the improvement of the effectiveness of analytical procedures has generally followed two approaches: (1) the use of sophisticated statistical techniques, and (2) the use of nonfinancial information.

These methods used in analytical procedures range from simple comparison (e.g. simple ratio analysis) to sophisticated statistical techniques (e.g. Vector Autoregression). Previous studies employ advanced statistical techniques to improve the accuracy and precision of the predictions of analytical procedures. For example, Dugan et al. (1985) introduce the Census X-11 Model as an improvement over the time-series model (ARIMA) by Kinney (1978) for analytical procedures. The X-11 time-series model decomposes the ARIMA model into three user-friendly components: trend-cycle, seasonal, and irregular movement. With the use of the three components, the X-11 model is validated to better capture the features of financial data than the ARIMA model. Wheeler and Pany (1990) compare the X-11 models with other alternatives, like regression models, Martingale models, Submartingale models in the performance of prediction and error detection. Their results indicate that the X-11 outperforms other models. Wild (1987) introduces a structural model that incorporates the components of both earnings and financial position into analytical procedures. Chen and Leitch (1998) extend the work of both Wheeler and Pany (1990) and Wild (1987) to develop a structural model that incorporates interdependency among the accounting data and exogenous variables. Their results indicate that the structural model dominates the prediction performance when it is compared to the ARIMA, the X-11, and Martingale models. Kogan et al. (2010) introduce continuous structural equations to analytical procedures for continuous auditing systems. Chen and Leitch (1999) show that the stepwise regression model is superior to the X-11, ARIMA, and Martingale models in discriminating between the decision risk and material error. Meanwhile, Dzeng (1994) introduces a new forecasting model called VAR (vector autoregression) to analytical procedures, and compares the model with seven other alternatives: three ARIMA models, two regression models, and two random walk models.

Early research indicates that auditors rely more heavily on financial information than nonfinancial information in establishing an overall level of audit scope, and consider nonfinancial information as corroborating evidence (Cohen, Krishnamoorthy, and Wright 2000). Trompeter and Wright (2010) show that auditors have recently gathered and considered more nonfinancial information that is available through the Internet. The increasing consideration and usage of nonfinancial information indicate the importance of nonfinancial information in analytical procedures.

Prior studies indicate that nonfinancial information can improve the accuracy of predictions in analytical procedures. Lev (1980) applies the economic and industrial data to analytical procedures and finds that it improves the forecasting ability of models. Brazel, Jones, and Zimbelman (2009) assess the inconsistent patterns between financial measures and nonfinancial measures and find that nonfinancial information can be used to detect financial statement fraud. A follow-up study by Brazel, Jones, Prawitt (2014) indicates that auditors generally do not react to the existence of the inconsistency between financial information and nonfinancial information, until a prompt calls auditors' attention to such inconsistency. Hoitash, Kogan, Vasarhelyi (2006) indicate that the contemporaneous financial information from peer companies in similar industries can improve the performance of analytical procedures. Previous work by Ittner and Larcker (1998) indicate that the customer satisfaction is partially reflected in current accounting numbers, and is correlated with future accounting numbers.

Search Volume

Researchers have used the SVI in various areas, including health phenomena, economics, finance, accounting, and marketing. Ginsberg et al. (2009) find that the SVI can predict flu outbreaks 1 to 2 weeks earlier than the US Center for Disease Control and Prevention (CDS) that rely on both virological test results and clinical data.

The extant research has examined the application of the SVI in the analysis and forecast of economic indicators. For example, Askitas and Zimmerman (2009) indicate the strong correlation between the SVI and unemployment rates in Germany. Vosen and Schmidt (2011) construct an indicator for private consumption using the SVI, and find that the indicator outperforms the existing survey-based indicators. Choi and Varian (2011) demonstrate how to use the SVI to predict economic indicators, including automobile sales, unemployment claims, consumer confidence, and travel destination planning. Wu and Brynjolfsson (2013) show that the model incorporating the SVI predicts housing market sales and prices more accurately than both the benchmark model and the predictions generated by experts from the National Association of Realtors. Li et al. (2015) apply the SVI to analyze trader positions and energy price volatility.

The extant literature indicates that the SVI can be a valid and direct proxy for investor attention (Da, Engelberg, and Gao 2011; Bank, Larch, and Peter 2011; Andrei and Hasler 2014; deHaan, Shevlin, and Thornock 2015; Goddard, Kita, and Wang 2015). For example, Da, Engelberg, and Gao (2011) introduces the SVI as a direct measure of investor attention and find that the SVI captures investor attention timelier than existing proxies, such as news coverage. Andrei and Hasler (2014) use the SVI as a proxy for investor attention to validate their theoretical expectation about the joint role played by investor attention and learning uncertainty on determining asset prices. deHaan, Shevlin, and Thornock (2015) employ the abnormal SVI as one of several empirical proxies for temporal variation in market attention to examine whether managers strategically release bad earnings news during periods of low market attention. Several studies have indicated that the SVI can perform well in terms of forecasting stock market activity (Siganos 2013; Dimpfl and Jank 2015).

Researchers have also suggested that the SVI reported by Google can be used as a good measure of information demand of retail investors (Vlastakis and Markellos 2012; Drake, Roulstone and Thornock 2012), and investor sentiment (Da, Engelberg, and Gao 2015; Joseph, Wintoki, and Zhang 2011). For example, Da, Engelberg, and Gao (2015) construct investor sentiment of U.S. households using the SVI, and further quantify the effects of investor sentiment on asset prices, volatility, and fund flows.

Many studies have considered the search volume from the perspective of individual investors. Few studies have considered the search volume from the perspective of consumers. Kulkarni, Kannan, Moe (2012) argue that the search volume provides valuable measures and indicators of consumer interest in products and develop a model using it to forecast new product sales. Du and Kamakura (2012) demonstrate the application of the search volume as an important marketing intelligence tool to identify and track the general tendencies in consumer interest and behavior.

The search volume has been used in various areas including health, economics, finance, accounting, and marketing. However, no study has applied it to the auditing domain. This study fills this gap by applying it to analytical procedures.

Research Questions

This study examines three research questions related to the usage of the consumer search volume in analytical procedures. The first research question examines whether the models incorporating the consumer search volume can generate more accurate and precise prediction than the models that do not use it.

The consumer search volume represents the general level of consumer interest in terms of products or services (Du and Kamakura 2012; Kulkarni, Kannan, Moe 2012). Consumers use search engines to seek products' information including reviews, deals, and so on. The search behavior of consumers indicates their interest in the products or brands. The search volume captures the interest of millions of customers over time. This proxy for consumer interest can be correlated with retail sales. The decrease of customer interest may lead to a decline of retail sales. The quantitative feature of the consumer search volume allows auditors to examine the trend and pattern of consumer interest. Therefore, auditors may benefit from examining the trend and pattern in the consumer search volume to gain insight into corporate performance.

The consumer search volume is believed to capture contemporaneous consumer interest. The consumer interest is contemporaneously captured through measuring consumers search behavior. The corporate sales may typically experience similar changes as indicated by customer interest. Thus, the consumer search volume is particularly well suited for predicting the contemporaneous corporate financial statement data. The prediction performance is evaluated based on the criterion of minimizing the mean absolute error percentage error (MAPE). The first question is shown below:

Research Question 1: Do the models that incorporate the consumer search volume generate smaller MAPE than the models that do not incorporate it?

The second research question examines the ability of the consumer search volume to detect errors when uncoordinated errors occur. Uncoordinated errors occur if errors exist in the dependent variable (e.g. overstated total revenue). The prior study indicates that the inconsistency between financial information and nonfinancial information will be a good indicator for firms' frauds or errors (Brazel, Jones, and Zimbelman 2009). This study expects that the consumer search volume working as nonfinancial information can be useful for fraud or error detection. For example, if the consumer search volume of a certain product online decreases but sales of the product increases, auditors could see this inconsistency as a potential red flag in terms of financial statement fraud or errors. Further, nonfinancial information is believed to be less likely to be manipulated by managers. The consumer search volume reported by Google is less likely to be manipulated as well. The search volume reported by Google does not count the repeated search behavior. Google normalizes the aggregative search volume to provide a search volume index for each term. The second research question is shown below:

Research Question 2: Are the models that incorporate the consumer search volume able to better detect uncoordinated errors than the models that do not incorporate it?

The third research question examines the ability of the model to detect errors when coordinated errors occur. Coordinated errors happen if errors exist both in the dependent accounts and the other related accounts (e.g. overstated total revenue and overstated accounts receivable). When performing analytical procedures, auditors may use the other related accounts to generate expectations. In the process of generating predictions, auditors generally assume the other related accounts are free from error.

However, managers may manipulate the other related accounts to be consistent with the errors in the dependent account. Thus, the auditors' analysis and judgment may be biased. The coordinated error is hard to detect since errors in accounts are in the same direction (i.e. both total revenue and accounts receivable are overstated). The third research question examines whether the models that incorporate the consumer search volume can better detect the coordinated errors than the models that do not incorporate it.

Research Question 3: Are the models that incorporate the consumer search volume able to better detect coordinated errors than the models that do not incorporate it?

2.3. Data and Methodology

Consumer Search Volume

The section begins by discussing the customer search volume, which is the main variable in this analysis. Google collects the search data of its search engine users and provides the monthly or weekly SVI of search terms through its product, Google Trends. Instead of providing the absolute search volume numbers, Google Trends indexed and normalized the search volume numbers through scaling them by the time-series average of the search window to make comparisons between terms easier. Google Trends returns a zero value for search terms if the search volume is very low during the week or month. Google Trends returns an empty SVI file if not enough search queries are submitted during the whole requested window.

One empirical concern is the identification of search terms in Google. A search term may have multiple meanings (e.g., "Apple", "Blackberry"). To solve the issue, Google Trends provides categories to refine the results across industries and topics.

Google Trends defines 25 main categories, and each main category has multiple subcategories. Queries on Google's search engine are automatically assigned to particular categories using natural language processing method. For example, the search of "McDonald's" can be limited to the "Food and Beverage" main category, which can be narrowed down to the "Restaurant" subcategory, or even further refined to the "Fast Food" sub-subcategory. Each category has a unique category code assigned by Google Trends. For example, the code of the "Food and Beverage" category code is "0-71", the code of its subcategory "Restaurant" is "0-71-276", and the code of the sub-subcategory "Fast Food" is "0-71-276-918". This study uses the category codes to refine the search results.

This study is cautious about using search terms in Google Trends. First, although the SVIs are refined with the categories, some terms are still generic under the specific category. Thus, the firms with generic terms are excluded from the sample. Second, Google Trends does not return the relative results of searches in different languages for the same terms. Customers in non-English speaking countries may use search terms in other languages. This leads to the failure of the search volume representing the general population. To alleviate the issue, this study only includes the firms that are incorporated in the United States of America.

This study builds a custom list of search terms for each firm. To capture consumer interest of each firm, the study employs corporate brands of each firm as search terms. The use of corporate brands as search terms is likely to represent the consumer search behavior related to its firm. A firm may have a large number of products (e.g. apparel industry) or have few products (e.g. automotive industry). Consumers may directly search brands to find products. Consumers who search for a certain product are likely to include the brand of the product into search queries to refine search results. Thus, corporate brands as search terms can capture the general level of consumer interest in terms of firms. To collect corporate brands, the study manually goes through corporate websites to find corporate brands. This has been done for all the firms in the sample. For firms having multiple brands, this study aggregates the SVI of each brand under the firms to represent consumer interest.

To collect the SVI data, the study employs a web crawling program that inputs the corporate brands and the category codes to download the SVI data into a CSV file. The filter is set for "2009-present". If Google Trends returns the weekly SVI, the weekly SVI will be further aggregated by month as the monthly SVI in order to merge with the monthly financial data. This has been done for all the firms in the sample. If corporate brands are rarely searched, Google Trends will return an empty file. For the firms with empty SVI files, this study removes them from the sample.

Sample Selection and Calculation of Variables

Corporate financial information is obtained from Standard and Poor's COMPUSTAT. This study extracts quarterly firm-specific variables for the period 2009 to 2015, especially information on total revenue, cost of goods sold, accounts receivable, accounts payable, and market sizes of firms.

For the purposes of the study, 23 customer-based industries were selected to construct the sample. The selected industry sectors include retail trade industries, food sector, apparel sector under manufacturing industry, accommodation sector, and food service industry. All the firms under the selected industrial sectors are initially included to avoid the subjectivity of sample construction. To remain in the sample, firms need to satisfy the following four requirements. First, firms need to be American companies to alleviate the impact of language. This is because Google Trend does not provide the relative results of searches in different languages for the same terms. Second, firms need to manufacture or sale products to customers. Or firms need to provide services directly to customers. Third, firms do not go bankrupt or are merged by other firms during the period 2009 to 2015. There are 465 firms satisfying the first three requirements. Fourth, firms need to have quarterly financial statement data for at least three years as well as non-zero SVI data for at least five years. There are 384 firms satisfying all four requirements. Table 2.1 presents the selected industry sectors and their North American Industry Classification System (NAICS) code. The final sample has 10,522 firms-quarterly observations, and the out-of-sample forecasting observations are 3,702. The selected industries are presented together with the industrial average revenue, cost of goods sold, accounts receivable, and accounts payable in Table 2.1.

Previous studies show that the disaggregated high frequency monthly financial data perform better than quarterly financial data in analytical procedures (Wild 1987; Dzeng 1994; Chen and Leitch 1998, 1999; Leitch and Chen 2003; Hoitash, Kogan, and Vasarhelyi 2006). Thus, this study uses the monthly financial data. Since firms only provide quarterly financial variables, this study uses the cubic splines method to interpolate monthly observations from quarterly observations. Monthly observations are generated for each of the four quarterly observations in a year. Leitch and Chen (1999) use the cubic splines interpolation method to interpolate quarterly financial data into monthly financial data, and claim that the method is the most frequent and best function to be employed for a curve fitting. Hoitash et al (2006) and Leitch and Chen (2003) also use this method to interpolate quarterly financial data into monthly data.

Table 2.1: Descriptive Statistics

NAICS Code	Industry Name	Number of Firm-Quarter	Total Revenue	Cost of Goods Sold	Accounts Receivable	Accounts Payable
311	Food Manufacturing	726	2527.67	1612.29	974.37	1009.99
312	Beverage and Tobacco Product Manufacturing	177	303.53	158.77	171.8	81.44
315	Apparel Manufacturing	567	716.12	366.95	311.7	165.54
316	Leather and Allied Product Manufacturing	405	765.51	387.96	370.4	197.24
325	Chemical Manufacturing	241	3251.65	983.8	1896.86	1158.45
334	Computer and Electronic Product Manufacturing	1091	3041.57	1779.4	1719.58	1314.76
336	Transportation Equipment Manufacturing	442	22198.28	17474.4	24933.46	10525.6
441	Motor Vehicle and Parts Dealers	517	1434.94	1078.73	196.96	486.57
442	Furniture and Home Furnishings Stores	215	690.4	409.31	22.27	229.08
443	Electronics and Appliance Stores	181	2704.11	1985.61	454.54	1312.61
444	Building Material and Garden Equipment and Supplies Dealers	171	6277.79	4089.35	404.88	2446.68
445	Food and Beverage Stores	323	4305.68	3270.19	260.39	909.61
446	Health and Personal Care Stores	376	6291.22	4969.89	1358.71	1651.92
447	Gasoline Stations	100	2377.27	2299.57	119.16	236.72
448	Clothing and Clothing Accessories Stores	1178	819.53	502.77	95.71	245.85
451	Sporting Goods, Hobby, Book, and Music Stores	312	892.24	576.77	356.27	419.06
452	General Merchandise Stores	475	12810.29	9412.9	843.01	4294.37
453	Miscellaneous Store Retailers	223	1512.56	1071.52	440.79	532.89
454	Nonstore Retailers	360	1647.69	1169.44	179.56	849.06
481	Air Transportation	484	2518.86	1961.16	448.68	507.63
492	Couriers and Messengers	63	6840.8	5877.53	2949.55	1021.53
721	Accommodation	477	1045.96	718.16	377.66	162.4
722	Food Services and Drinking Places	1418	582.38	423.1	78.53	96.22

This table presents descriptive statistics for the 23 selected industries between the years 2009-2015. The number of quarterly observations in the sample is presented for each industry. The average of the accounts balance for the total sales, cost of goods sold, accounts receivable, accounts payable are broken by the first three-digit NAICS code.

This study employs economic and industrial variables that were obtained from publicly available sources. The macroeconomic data, GDP, are collected from U.S. Department of Commerce Bureau of Economic Analysis[2]. This study uses the current-dollar seasonally adjusted quarterly GDP rates. The industrial data are calculated using the quarterly financial information of all the firms under the same first three-digit NAICS code. To match with monthly financial observations, the quarterly GDP rates and the industrial average are further interpolated into monthly observations using the cubic splines method.

To evaluate the error detection ability of the consumer search volume, this study simulates errors and seed them into account balances. Specifically, this study seeds errors into actual revenue to evaluate the ability of the model that incorporates the consumer search volume in detecting the existence of the errors. Two types of errors are simulated: uncoordinated errors and coordinated errors. Uncoordinated errors are the errors that only exist in the predicted variable, specifically overstated total revenue. Coordinated errors are the errors that exist both in the predicted accounts and in the other related accounts. This makes coordinated errors hard to be detected because of the same direction errors existing in the other related accounts. Specifically, the coordinated errors are overstated total revenue and overstated accounts receivable. The error rates seeded into accounts are 20% of the accounts balance.

An external auditor requires an investigation when the difference between the actual account balance and the predicted account balance exceeds the auditor's specific risk level. Using statistical methods, an investigation takes place when the standardized

[2] http://www.bea.gov/national/

difference between the actual and the predicted accounts balance exceeds the Z-value of the specific risk level that was set by auditors. In this experiment, the false negative error occurs when an error was seeded into the accounts balance but the statistical investigation fails to raise the alarm. The false positive error occurs when no error was seeded into the account balances but the statistical investigation rule raises an alarm. This study uses the false negative rates and the false positive rates to evaluate the performance of the models in error detection. The lower the false negative rates and the false positive rates are, the better the models perform.

Expectation Models

This study generates the expectation model using the customer search volume to predict the monthly accounts balance in analytical procedures. The expectation model is constructed to include prior financial information, economic and industrial indicators, and the consumer search volume represented. These variables are further explained below. First, financial information of the current period is highly correlated with financial information of its comparable prior period. Thus, financial information of comparable prior periods is commonly included in analytical procedures to generate expectation for the current year's financial performance (SAS No.56; SAS No. 122; Hoitash et al. 2006). Second, firms are expected to share the same industrial and economic effect even though they are different from locations or business. Previous studies have validated the contribution of economic and industrial indicators to the effectiveness of analytical procedures (Lev 1980; Loebbecke and Stenbart 1987; Wild 1987). Third, the SVI reported by Google is included to capture contemporaneous consumer interest. The SVI allows auditors to analyze the trends and change of consumer interest and to gain insight into the current period's financial performance.

The expectation model that incorporates the consumer search volume is compared with the benchmark model on the effectiveness of generating predictions for analytical procedures. The benchmark models are derived using prior financial information, and economic and industrial indicators. This study uses the ordinary least squares regression model to generate the expectation models. Previous studies show that regression models are more efficient than traditional procedures, and offer considerable improvement in predictions over those of naïve models, or non-statistical models (Kinney, 1978; Wilson, 1991). The specification of the expectation models are shown below:

$$X_{jt} = \beta_0 + \beta_1 X_{jt-12} + \beta_2 GDP_t + \beta_3 Industry_t + \varepsilon \qquad (2.1)$$

$$X_{jt} = \beta_0 + \beta_1 X_{jt-12} + \beta_2 GDP_t + \beta_3 Industry_t + \beta_4 SVI_{jt} + \varepsilon \qquad (2.2)$$

X_{jt} in the models is either total revenue or cost of goods sold (COGS) of firm *j* in month *t*. X_{jt-12} is either total revenue or COGS of the same month in the last year. SVI_{jt} is the consumer search volume index of firm j in month *t*. *GDP* is in nominal value. *Industry* is the industrial average, which is constructed using all the firms under the same three-digit NAICS code. Two different models are examined for each firm, respectively. Model 2.1 is the benchmark model. Model 2.2 is the model incorporating the consumer search volume.

Contemporaneous other related accounts have the potential contribution to the prediction accuracy of the predicted accounts (Hoitash et al. 2006). Specifically the contemporaneous accounts receivable may improve the prediction of total revenue, and the contemporaneous accounts payable may improve the accuracy of COGS. The study includes the contemporaneous other related accounts into the expectation models to see whether the model incorporating the consumer search volume can still improve the accuracy of the prediction.

$$X_{jt} = \beta_0 + \beta_1 X_{jt-12} + \beta_2 GDP_t + \beta_3 Industry_t + \beta_4 Account_{jt} + \varepsilon \quad (2.3)$$

$$X_{jt} = \beta_0 + \beta_1 X_{jt-12} + \beta_2 GDP_t + \beta_3 Industry_t + \beta_4 Account_{jt} + \beta_5 SVI_{jt} + \varepsilon \quad (2.4)$$

Account is accounts receivable if the independent variable is total revenue, and is accounts payable if the independent variable is COGS. Other variables are defined the same as before. Model 2.3 and Model 2.4 are the expectation model adding contemporaneous other related accounts. Model 2.4 is the expectation model further incorporating the consumer search volume. Each model (Model 2.1-2.4) is separately estimated for each firm for the period 2009-2014, and then the out-of-sample forecasts are generated for each firm for 2015.

The prediction performance is evaluated based on the criterion of minimizing the mean absolute error percentage error (MAPE)[3]. The MAPE is calculated for each firm using the out-of-sample prediction error of each company-month observation. Then the MAPE is aggregated over industry or firm size to see the prediction performance of the models by industry or by the market size of firms.

2.4. Results

Prediction Performance

[3] $MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|Actual\ Value_t - Forecast\ Value_t|}{Accutal\ Value_t}$

The first research question examines whether the model incorporating the consumer search volume provides more accurate prediction than the benchmark model in analytical procedures. This study calls the model that incorporates the consumer search volume the SVI model. This study evaluates the performance of the models on the prediction of the total revenue using Models 2.1 and 2.2. The MAPE was calculated for each firm of the out-of-sample forecasts. The MAPE are further aggregated by industry to compare the prediction performance of the SVI model with the benchmark model by industry. Table 2.2 presents the aggregative MAPE for each one of the 23 selected industries for the SVI model and the benchmark model. Two-tailed t-tests were examined to assess a statistical difference in the MAPE between the SVI model and the benchmark model in each industry. As can be seen, the MAPE of the SVI model is generally smaller than the MAPE of the benchmark model. The results of t-test show that the SVI model statistically outperforms the benchmark models for 11 industries. Thus the result indicates that the model with the consumer search volume improves the prediction accuracy for 11 industries.

The study further examines the prediction performance of the models when contemporaneous accounts receivable is added. Models 2.3 and 2.4 are separately evaluated for each firm in the sample. Table 2.3 displays the aggregative MAPE by industry and the t-test results. As displayed in Table 2.3, the contemporaneous firm-specific data improve the performance of the models on the out-of-sample forecasts. The SVI model is statistically superior to the benchmark model for 18 industries on the out-of-sample forecasts. The SVI model improves the prediction of total revenue for food manufacturing, leather and allied product manufacturing, motor vehicle and parts dealers, electronics and appliance stores, food and beverage stores, health and personal care stores, clothing and clothing accessories stores, air transportation, and so on. The benchmark model outperforms the SVI model for transportation equipment manufacturing, furniture and home furnishings stores, Building Material and Garden Equipment and Supplies Dealers, Gasoline Stations, General Merchandise Stores.

Table 2.2: Aggregative Prediction Performance for Total Revenue by Industry (Model 2.1 and Model 2.2)

NAICS Code	Number of Firm-Months	Benchmark-model MAPE	SVI Model MAPE	T-test
311	267	0.1175	0.1136	0.969**
312	63	0.0928	0.0917	0.562
315	189	0.119	0.1092	1.000***
316	150	0.1328	0.1176	1.000***
325	90	0.2093	0.2016	0.984**
334	327	0.4574	0.3882	0.867
336	147	0.1844	0.1854	0.393
441	192	0.12	0.116	0.958**
442	96	0.1094	0.1085	0.6
443	72	0.0974	0.0937	0.943*
444	54	0.1567	0.1573	0.43
445	111	0.08	0.0759	0.990***
446	141	0.197	0.1877	0.985**
447	45	0.2174	0.2152	0.719
448	414	0.0955	0.091	1.000***
451	105	0.0884	0.087	0.817
452	174	0.0911	0.0925	0.022
453	78	0.1708	0.1721	0.212
454	147	0.3438	0.372	0.052
481	147	0.1156	0.101	1.000***
492	12	0.1901	0.1902	0.484
721	177	0.2317	0.2349	0.326
722	504	0.1167	0.1085	0.999***

This table presents the prediction results of Model 2.1 and Model 2.2 for total revenue. The table shows the MAPE of the out-of-sample forecasts of both models by industry. T-test was examined to assess the statistical difference of the MAPEs of the models. ***,**,* represent significance beyond the 99th, 95th, and 90th percentile levels, respectively.

Table 2.3: Aggregative Prediction Performance for Total Revenue by Industry (Model 2.3 and Model 2.4)

NAICS Code	Number of Firm-Months	Benchmark-model MAPE	SVI Model MAPE	T-test
311	276	0.0988	0.0939	0.996***
312	63	0.118	0.1018	0.993***
315	204	0.1024	0.0922	1.000***
316	150	0.0826	0.0678	1.000***
325	90	0.2022	0.1949	0.996***
334	387	0.2875	0.2327	0.950**
336	159	0.1292	0.1293	0.487
441	192	0.1012	0.0988	0.996***
442	96	0.097	0.0966	0.555
443	72	0.1212	0.1135	0.999***
444	54	0.1346	0.1351	0.408
445	120	0.0788	0.0739	0.974**
446	141	0.1449	0.1327	1.000***
447	45	0.1882	0.1883	0.478
448	414	0.0903	0.0873	1.000***
451	117	0.0791	0.077	0.924*
452	174	0.092	0.0938	0.009
453	78	0.2158	0.2138	0.979**
454	135	0.4046	0.3803	0.988**
481	171	0.0939	0.0911	0.966**
492	21	0.0565	0.0518	0.970**
721	195	0.186	0.1805	0.998***
722	552	0.0951	0.0918	0.993***

This table presents the prediction results of Model 3 and Model 4 for total revenue. The table shows the MAPE of the out-of-sample forecasts of both models by industry. T-test was examined to assess the statistical difference of the MAPEs of the models. ***,**,* represent significance beyond the 99th, 95th, and 90th percentile levels, respectively.

Table 2.4 displays the forecasting results of Model 2.3 and 2.4 by firms' market size to provide insight into the improvement of the customer search volume on analytical procedures. The market size of firms is obtained from COMPUSTAT. The market size of each firm is logged (base ten) prior to being rounded to the nearest integer.

Firms in the sample are broken down by the market size 0-13. Table 2.4 presents the aggregative MAPE by size, the results of t-test. The results are broken down by firm size. The MAPE and the t-test results indicate that the SVI model is statistically superior to the benchmark model when the market size of firms ranges from 3 to 10. When the market size of firms is small (0-2) or large (11-13), the benchmark model outperforms the SVI model. This may be explained by the following reasons. Google Trends returns a zero value of the SVI when the submitted search volume of the term is low. Small-sized firms are generally less exposed to the public and may have low search traffic that leads to zero values of the SVI. Large firms are more likely to be multinational companies and sell products worldwide. The multinational companies selling products in non-English speaking countries may use the local language to name corporate brands. To date, Google Trends has not returned the relative results of searches in different languages for the same terms. Thus, the SVI may not represent the general population of consumers search behavior. As a result, the model incorporating the consumer search volume does not improve the performance of analytical procedures for large size firms.

Table 2.4: Aggregative Prediction Performance for Total Revenue by Size (Model 2.3 and Model 2.4)

Market Size	Number of Firm-Months	Benchmark-model MAPE	SVI model MAPE	T-test
1	18	0.3052	0.3039	0.55
2	33	1.1493	0.7148	0.914*
3	66	0.1843	0.1773	0.994***
4	171	0.1309	0.1272	0.999***
5	300	0.1519	0.1413	1.000***
6	435	0.1038	0.0994	1.000***
7	513	0.1864	0.1781	0.996***
8	681	0.1068	0.0992	1.000***
9	477	0.1225	0.1151	1.000***
10	426	0.106	0.1018	1.000***
11	180	0.121	0.1216	0.375
12	75	0.0631	0.0617	0.722
13	27	0.115	0.1141	0.621

This table presents the prediction results of Model 3 and Model 4 for total revenue by the market size of firms. The table shows the MAPE of the out-of-sample forecasts of both models by the market size of firms. T-test was examined to assess the statistical difference of the MAPEs of the models. ***,**,* represent significance beyond the 99th, 95th, and 90th percentile levels, respectively.

Error Detection

The third research question examines the error detection performance of the SVI model using simulated errors. The SVI model captures the consumer interest in corporate products and services. The inconsistency between consumer interest and financial statement information may be a signal for firms' potential errors. The SVI model is expected to improve the error detection ability. To examine this, this study seeds errors into total revenue to see whether the SVI model outperforms the benchmark model in detecting errors. Table 2.5 presents the performance of the expectation models on error detection. For 17 industries, the SVI model results in both less false negative and false positive errors than the benchmark model. The SVI model performs similar with the benchmark model for two industries. For three industries, the SVI model leads to more false negative errors and less false positive errors than the benchmark model.

Further, the study examines the error detection ability of the SVI model when coordinated errors occur. In the coordinated error scenario, both accounts receivable and total revenue are overstated through seeding with the same rate of errors. Table 2.6 presents the percentage of false positive and false negative errors for both the SVI model and the benchmark model. The SVI model leads to both less false negative and false positive errors than the benchmark model for 14 industries. The benchmark model outperforms the SVI model in detecting coordinated error for one industry. The results are mixed for seven industries.

Table 2.5: False Positive and False Negative Errors for Uncoordinated Error

NAICS Code	Number of Firm-Months	Benchmark Model		SVI Model		Improved False Negative	Improved False Positive	Superior Model
		False Negative	False Positive	False Negative	False Positive			
311	276	37.68%	21.74%	37.32%	19.20%	0.36%	2.54%	SVI model
312	63	68.25%	20.63%	63.49%	11.11%	4.76%	9.52%	SVI model
315	204	55.88%	16.18%	51.96%	11.27%	3.92%	4.90%	SVI model
316	150	57.33%	8.67%	50.00%	4.67%	7.33%	4.00%	SVI model
325	90	45.56%	40.00%	45.56%	38.89%	0.00%	1.11%	SVI model
334	387	48.32%	17.05%	46.25%	16.80%	2.07%	0.26%	SVI model
336	159	38.99%	23.90%	35.22%	21.38%	3.77%	2.52%	SVI model
441	192	48.44%	18.75%	48.44%	17.71%	0.00%	1.04%	SVI model
442	96	42.71%	13.54%	40.63%	13.54%	2.08%	0.00%	SVI model
443	72	58.33%	6.94%	55.56%	4.17%	2.78%	2.78%	SVI model
444	54	35.19%	16.67%	35.19%	20.37%	0.00%	-3.70%	Benchmark model
445	120	30.83%	7.50%	33.33%	5.83%	-2.50%	1.67%	Mixed
446	141	26.24%	31.21%	21.99%	27.66%	4.26%	3.55%	SVI model
447	45	84.44%	46.67%	84.44%	46.67%	0.00%	0.00%	-
448	414	37.68%	15.70%	35.75%	14.49%	1.93%	1.21%	SVI model
451	117	42.74%	6.84%	40.17%	4.27%	2.56%	2.56%	SVI model
452	174	33.91%	13.22%	32.76%	13.22%	1.15%	0.00%	SVI model
453	78	29.49%	25.64%	29.49%	25.64%	0.00%	0.00%	-
454	135	59.26%	24.44%	60.74%	23.70%	-1.48%	0.74%	Mixed
481	171	16.96%	5.26%	18.13%	4.09%	-1.17%	1.17%	Mixed
492	21	9.52%	14.29%	9.52%	9.52%	0.00%	4.76%	SVI model
721	195	47.18%	23.08%	46.67%	22.56%	0.51%	0.51%	SVI model
722	552	25.91%	12.86%	24.82%	12.14%	1.09%	0.72%	SVI model

This table presents the error detection performance of Model 2.3 and Model 2.4 after errors are seeded into total revenue (i.e., uncoordinated errors). Errors are seeded separately into each monthly financial observation of each firm. Statistical investigation rule applies to detect these errors. The auditors' specific risk level takes 0.05. The seeded error rate is 0.2. The False Positive and False Negative rates are broken down by industry. The improved false positive or false negative errors are calculated by the difference between the SVI model (Model 2.4) and the benchmark model (Model 2.3). The superior model is determined based on the sign of improved false positive and improved false negative errors .

Table 2.6: False Positive and False Negative Errors for Coordinated Error

NAICS Code	Number of Firm-Months	Benchmark Model		SVI Model		Improved False Negative	Improved False Positive	Superior Model
		False Negative	False Positive	False Negative	False Positive			
311	276	65.94%	35.51%	68.84%	34.42%	-2.90%	1.09%	Mixed
312	63	82.54%	31.75%	82.54%	28.57%	0.00%	3.17%	SVI Model
315	204	73.04%	30.88%	74.02%	25.49%	-0.98%	5.39%	Mixed
316	150	80.00%	28.00%	79.33%	22.00%	0.67%	6.00%	SVI Model
325	90	66.67%	46.67%	67.78%	42.22%	-1.11%	4.44%	Mixed
334	387	61.50%	29.20%	61.76%	26.36%	-0.26%	2.84%	Mixed
336	159	49.69%	38.99%	46.54%	38.36%	3.14%	0.63%	SVI Model
441	192	59.38%	29.69%	58.33%	27.08%	1.04%	2.60%	SVI Model
442	96	47.92%	12.50%	43.75%	11.46%	4.17%	1.04%	SVI Model
443	72	63.89%	4.17%	62.50%	4.17%	1.39%	0.00%	SVI Model
444	54	64.81%	37.04%	64.81%	37.04%	0.00%	0.00%	-
445	120	39.17%	13.33%	40.00%	11.67%	-0.83%	1.67%	Mixed
446	141	39.01%	34.04%	29.79%	32.62%	9.22%	1.42%	SVI Model
447	45	80.00%	46.67%	77.78%	46.67%	2.22%	0.00%	SVI Model
448	414	43.00%	16.18%	40.58%	15.46%	2.42%	0.72%	SVI Model
451	117	46.15%	5.98%	45.30%	5.13%	0.85%	0.85%	SVI Model
452	174	36.78%	11.49%	35.06%	12.07%	1.72%	-0.57%	Mixed
453	78	48.72%	33.33%	47.44%	33.33%	1.28%	0.00%	SVI Model
454	135	68.15%	35.56%	68.15%	33.33%	0.00%	2.22%	SVI Model
481	171	35.67%	1.17%	35.67%	2.34%	0.00%	-1.17%	Mixed
492	21	14.29%	23.81%	19.05%	28.57%	-4.76%	-4.76%	Benchmark model
721	195	46.67%	31.28%	46.67%	29.74%	0.00%	1.54%	SVI Model
722	552	32.97%	18.48%	32.43%	15.58%	0.54%	2.90%	SVI Model

This table presents the error detection performance of Model 2.3 and Model 2.4 after errors are seeded into total revenue and accounts receivable (i.e., coordinated errors). Errors are seeded separately into each monthly financial observation of each firm. Statistical investigation rule applies to detect these errors. The auditors' specific risk level takes 0.05. The seeded error rate is 0.2. The False Positive and False Negative errors are broken down by industry. The improved false positive or false negative errors are calculated by the difference between the SVI model (Model 3.4) and the benchmark model (Model 2.3). The superior model is determined based on the sign of improved false positive and improved false negative.

2.5. Additional Analysis

This study further examines whether the consumer search volume can be used as nonfinancial information to improve the prediction for the cost of goods sold (COGS). Firms need to match the COGS along with sales. For some industries such as apparel manufacturing, the COGS is added up with the units of product sold. But for some industries such as transportation equipment, the COGS is not consistent with the units of products sold. This study estimates Model 2.4 using the COGS and the accounts payable, and compares the estimated results to the benchmark model, Model 2.3. The t-test is used to estimate the difference of the MAPE in prediction by industry. Table 2.7 presents the aggregative MAPE by industry and the t-tests. The results show that the SVI model provides smaller MAPE than the benchmark model for 13 out of 23 industries. The results indicate that auditors can employ the consumer search volume to improve the prediction of COGS for some industries as well.

Table 2.7: Aggregative Prediction Performance for Cost of Goods Sold with AP by Industry (Model 2.3 and Model 2.4)

NAICS Code	Number of Firm-Months	Benchmark-model MAPE	SVI Model MAPE	T-test
311	261	0.1246	0.1268	0.21
312	63	0.1348	0.1144	1.00***
315	189	0.1483	0.14	1.00***
316	150	0.1484	0.1381	1.00***
325	87	0.3724	0.3747	0.42
334	318	0.2939	0.2358	0.98**
336	144	0.1618	0.163	0.35
441	189	0.0989	0.0973	0.86
442	96	0.0953	0.0918	0.84
443	72	0.1157	0.117	0.3
444	54	0.164	0.1618	0.88
445	111	0.0797	0.0757	0.99***
446	141	0.1587	0.1532	0.95**
447	45	0.1638	0.1635	0.57
448	414	0.1181	0.1139	1.00***
451	105	0.0844	0.0827	0.82
452	174	0.09	0.0918	0.06
453	78	0.1918	0.1946	0.12
454	144	0.238	0.2719	0.12
481	141	0.2069	0.197	1.00***
492	9	0.1434	0.1433	0.57
721	168	0.2535	0.2428	0.96**
722	477	0.1703	0.1744	0.11

This table presents the prediction results of Model 2.3 and Model 2.4 for cost of goods sold. The table shows the MAPE of the out-of-sample forecasts of both models by industry. T-test was examined to assess the statistical difference of the MAPEs of the models.

For robustness and sensitivity testing, the auditor risk level and the seeded error rates were set to different values. This study uses a simulated experiment to examine the ability of the SVI model on error detection. Simulated errors were seeded into the total revenue and accounts receivable. The statistical investigation rule is generated based on the auditors' specific risk level. The auditor's specific risk level is set to be

$\alpha = 0.1$ for the statistical investigation rule when seeded error rates are 0.2. Table 2.7 presents the false negative and false positive errors broken down by industry. The results show that the SVI model is superior to the benchmark model for 15 out of 23 selected industry.

Table 2.8: False Positive and False Negative Rates for Uncoordinated Error (Alpha=0.10)

NAICS Code	Number of Firm-Months	Benchmark Model		SVI Model		Improved False Negative	Improved False Positive	Superior Model
		False Negative	False Positive	False Negative	False Positive			
311	276	26.81%	33.33%	25.72%	30.80%	1.09%	2.54%	SVI model
312	63	58.73%	31.75%	50.79%	22.22%	7.94%	9.52%	SVI model
315	204	44.61%	25.00%	41.67%	21.08%	2.94%	3.92%	SVI model
316	150	46.00%	17.33%	38.00%	12.00%	8.00%	5.33%	SVI model
325	90	36.67%	48.89%	35.56%	46.67%	1.11%	2.22%	SVI model
334	387	36.43%	21.96%	36.43%	20.93%	0.00%	1.03%	SVI model
336	159	25.16%	29.56%	22.64%	27.04%	2.52%	2.52%	SVI model
441	192	30.21%	26.04%	33.33%	25.00%	-3.13%	1.04%	Mixed
442	96	30.21%	17.71%	23.96%	16.67%	6.25%	1.04%	SVI model
443	72	47.22%	15.28%	48.61%	13.89%	-1.39%	1.39%	Mixed
444	54	33.33%	24.07%	33.33%	24.07%	0.00%	0.00%	-
445	120	20.00%	20.00%	22.50%	18.33%	-2.50%	1.67%	Mixed
446	141	19.15%	38.30%	15.60%	34.75%	3.55%	3.55%	SVI model
447	45	82.22%	53.33%	80.00%	53.33%	2.22%	0.00%	SVI model
448	414	26.57%	22.95%	25.36%	21.26%	1.21%	1.69%	SVI model
451	117	30.77%	11.97%	29.91%	10.26%	0.85%	1.71%	SVI model
452	174	22.99%	16.67%	22.41%	17.82%	0.57%	-1.15%	Mixed
453	78	24.36%	32.05%	23.08%	33.33%	1.28%	-1.28%	Mixed
454	135	51.11%	32.59%	54.07%	31.11%	-2.96%	1.48%	Mixed
481	171	8.77%	9.36%	8.19%	8.19%	0.58%	1.17%	SVI model
492	21	9.52%	14.29%	9.52%	14.29%	0.00%	0.00%	-
721	195	38.46%	32.31%	36.41%	30.26%	2.05%	2.05%	SVI model
722	552	19.38%	17.93%	18.48%	17.57%	0.91%	0.36%	SVI model

This table presents the error detection performance of Model 2.3 and Model 2.4 after errors are seeded into total revenue (i.e., uncoordinated errors). Errors are seeded separately into each monthly financial observation of each firm. Statistical investigation rule applies to detect these errors. The auditors' specific risk level takes 0.10. The seeded error rate is 0.2. The False Positive and False Negative rates are broken down by industry. The improved false positive or false negative is calculated by the difference between the SVI model (Model 2.4) and the benchmark model (Model 2.3). The superior model is determined based on the sign of improved false positive and improved false negative.

2.6. Conclusion and Limitations

This study examines whether the consumer search volume can improve prediction performance and error detection in analytical procedures. The consumer search volume captures the general level of consumer interest in terms of corporate products or services. This information can be used by auditors to gauge sales and improve analytical procedures. The results indicate that the model with the consumer search volume generates smaller MAPE than the benchmark models for most consumer-based industries. A simulated experiment was conducted to examine the ability of the SVI model in error detection. The results indicate the SVI model generally improves the error detection when uncoordinated errors and coordinated errors occur.

This study is limited for the following reasons. First, this study does not differentiate the economic importance of consumers to firms. Some consumers may contribute more to the economic situation of the companies. The SVI only counts the amount of submitted queries from all consumers. Second, the consumer search volume measures consumer interest instead of actual consumer purchases. Third, this study assumes that the consumer search volume is positively related to consumer interest. Consumers may search firms due to concerns about some unrelated events. For example, a labor strike may attract the attention of consumers in a negative manner.

Chapter 3 Social Capital and the Municipal Bond Market

3.1. Introduction

It is well documented that the municipal bond market is subject to a much lesser degree of the Security and Exchange Commission's (SEC) reporting requirements than the corporate capital markets[1]. The municipal bond market is considered to be more opaque, resulting in higher uncertainty and information asymmetry between the bond issuers and the bondholders. This opaqueness increases the risk associated with municipal bonds. Additionally, the risk of municipal bonds is greater if the social environment of municipalities is associated with bribery, fraud, and other illegal activities[2,3]. The negative impact of the social environment is especially highlighted by the recent scandals, which contributed to the bankruptcies of municipalities such as Jefferson County [4], Detroit, Orange County, and several other Californian municipalities[5,6].

---

1 Unlike publicly held corporations that are all subject to the regulatory authority of the Securities and Exchange Commission, local governments face different state government regulations. Some states require GAAP compliance, some require compliance with state designated (non-GAAP) disclosure practices, and some do not regulate local government financial disclosures (Ingram and Dejong 1987 (Abstract)).

2 For example, see Mitchell and Vogel, "Illegal Payments Mar the Muni Market," The Wall street Journal (May 5, 1993)

3 Despite several regulatory attempts to address some of these concerns, certain negative aspects of the information environment continue to exist.

4 Jefferson County, New Jersey and the city of Birmingham are recent high profile examples of fraud, corruption and the failure to properly disclose information to investors. http://www.realclearmarkets.com/articles/2010/09/01/how_corrupt_is_the_muni-bond_business_98651.html.

5 Detroit's bankruptcy had a significant impact on the municipal bond market, and some municipalities postponed the offering on new issues due to this. http://www.thefiscaltimes.com/Columns/2013/08/07/6-Muni-Bond-Myths-Rocking-the-Market

6 Growing financial distress in many California cities put bondholders at risk. http://www.huffingtonpost.com/2012/08/17/california-bankruptcies_n_1799543.html

The opaqueness of the municipal bond market and the limited regulatory oversight is complicated by the fact that the municipal bond market currently has more than one and half million different types of bonds outstanding. As a result, investors of municipal bonds face substantial search costs in order to price individual credits. Given these characteristics, it is important to understand factors, which investors may rely on to price risk. The municipal bond market is also important to understand because it represents approximately $3.7 trillion[7] in outstanding debt. As a source of funding for essential services for state and local governments, the municipal bond market has grown in size from 185 billion in bond issuances in 1996 to over 433 billion in 2010.

This study examines whether the levels of social capital in municipalities influence risk associated with municipal bonds. Defined as the norms and networks that encourage cooperation, social capital can be thought of as a social construct, which captures a region's level of altruism, trustworthiness, and propensity to honor obligations (Putnam, 1995). Woodcock (2001) argues that social and ethical norms fostered by high levels of social capital could promote cooperation and a greater sense of trust. Studies also suggest that social capital encourages more honesty and less fraudulent behavior in bureaucrats (LaPorta et al., 1997; Buonanno et al., 2009). Consistent with this line of reasoning, Jha and Chen (2014) document that a greater level of trust among auditors and managers in high social capital areas results in lower audit fees.

Based on the existing evidence, this study argues that high social capital will create higher trust and cooperation among municipal bond issuers and investors. This will in turn result in higher reliability of information provided by issuers to the potential

[7] Sifma research, outstanding amount as of 2010.

bondholders, which will reduce risk and hence borrowing costs. Thus, this study examines the first research question on the association between social capital and bond costs, measured by bond yield, on a large sample of municipal bonds issued from 1998 through 2012. This time period is selected because it provides us with a unique opportunity to address the research question when two major bond insurers were downgraded in the aftermath of the financial crisis, which considerably increased the need for investors to monitor risk. This study also examines whether both general obligation bonds and revenue bonds will be influenced by social capital. This study expects that social capital will especially affect bond costs of general obligation bonds because general tax revenues support these bonds. On the other hand, revenue bonds are not covered by general tax revenues and are backed by project specific revenues.

The municipal bond market is unique in that a large portion of bonds is sold with default insurance provided by one of the major insurers such as AMBAC and MBIA. Bond insurance protects investors from bond default and also reduces the need for investors to monitor municipalities. The credit rating downgrades and loss of triple-A ratings for AMBAC and MBIA changed market dynamics. Reports show that the demand for insurance fell considerably after their initial downgrades in 2008[8]. This study uses this event to examine whether social capital is viewed as a substitute for insurance by investors. High social capital environments are associated with more efficient government, greater trustworthiness, and the propensity to honor obligations. Thus, high social capital may reduce the risk associated with municipal bonds. This also implies that high social capital reduces the need to monitor municipalities. This

[8] From as high as 57% insured before the crisis to 19% in 2008, and to 3.5% in 2012. http://www.bondbuyer.com/issues/123_83/bond-insurance-then-and-now-revival-of-industry-1062071-1.html.

study conducts two tests revolving around bond insurance in the municipal bond market. First, this study examines whether the demand for insurance is low for bonds issued in high social capital environments. This argument suggests that as a substitute, social capital may moderate the demand for insurance. Second, this study examines the association between social capital and bond yields in the pre and post insurer downgrade periods. Without high quality insurance that serves to monitor municipalities, it is expected that investors will rely more on social capital as a source of information. Using this exogenous shock, this study tests whether social capital is more associated with bond risk in the post downgrade period.

The test results show that there is a significantly negative association between bond yield and social capital, confirming that bond cost is lower for municipal bonds that are issued by the municipalities located in high social capital counties. The results also show a statistically negative association between yield spread and social capital, confirming that bond risk is high for bonds that are issued by the municipalities located in low social capital regions. These findings are interpreted to suggest that high social capital serves as an alternative source to evaluate the reliability of information of provided by municipalities.

The results also indicate that the general obligation bonds issued by municipalities located in the counties with high social capital are associated with lower bond yield and lower yield spread. On the other hand, the study finds no impact of social capital on the revenue bonds. The differential impact of social capital on general obligations and revenue bonds is probably due to the sources of funds that support the payment of interest and principal. The principal and interest of general obligation bonds are secured by local tax and the revenue bonds are secured by projects' revenue.[9]

The study evaluates the impact of Moody's downgrading of insurers AMBAC and MBIA by comparing the association of social capital with bond yield and yield spread for the pre-downgrade period with the post-downgrade period. The results show that the association between social capital and the cost of borrowing was greater after June 2008 when the insurers were downgraded from triple A to a lower rating. The results show that the economic impact of social capital on the yield spread is considerably greater in the post-period. The results show a movement from the 25th to 75th percentile of social capital with about a 7 percent decrease in bond yields and about a 5.7 percent decrease in yield spread. This decrease is similar to the impact from a change of one standard deviation in the issue size of a bond.

Additionally, the results show that the demand for insurance is high when the issuers are located in low social capital counties. The results show that the likelihood of insurance decreases 1.15 times for a one standard deviation change in the social capital. This finding supports the expectation that the purchase of insurance is negatively associated with social capital.

This study conducts additional tests to evaluate the robustness of the findings. First, the study examines the association between bond prices in the secondary market and social capital on the assumption that the secondary market prices will allow us to evaluate the association between risk and social capital irrespective of the decision to issue new debt. Consistent with the main results, the results indicate a greater

---

9 Alm and Gomez (2008) claims that social capital plays an important role in determining an individual's intrinsic motivation to pay tax.

association between social capital and bond prices in the secondary market after the downgrade of the bond insurers. Furthermore, the results indicate that the impact of social capital on trade price is statistically significant after the downgrade of insurers for insured general obligation bonds. Thus, this analysis also confirms that social capital becomes more important as a source of information in the absence of high quality insurers.

Second, the study conducts a test on a smaller sample of bonds from states with high variation in social capital. The results of these tests are consistent with the main tests. Thus, these results provide evidence that social capital plays an important role in providing information on the risk of the municipal bond market and the reliability of information provided by the bond issuers.

This study makes the following contributions to the literature. First, consistent with the economics literature that emphasizes the role of social capital in the production process (Fukuyama 1995; Guiso, Sapienza, and Zingales, 2000; Rupasingha, Goetz, and Freshwater, 2006), this study provides evidence that bond costs are also affected by the social capital. Second, the findings provide a better understanding of the determinants of borrowing for municipal financing. While the existing studies provide evidence on the competition among underwriters (Kessel, 1971), insurance (Kidwell, Sorensen, and Wachowicz, 1987), fiscal decisions (Capeci, 1994), disclosure regulation (Benson et al. 1991; Baber and Gore, 2008), and accounting restatements (Baber et al., 2013), the findings provide evidence that social capital also affects the borrowing costs. Third, consistent with the existing studies that provide evidence on the effect of social capital on financial quality and audit fees (Jha, 2013; Jha and Chen, 2014), the findings of this study show that social capital also has a significant impact on the risk environment of the municipal bond market.

The rest of the paper is organized as follows. Section 3.2 contains the municipal bond background, social capital and hypothesis. Section 3.3 shows the models and data. Results are shown in Section 3.4 and Section 3.5 concludes.

3.2. Background and Hypothesis Development

Municipal Bonds Market

Municipal bonds are debt securities that may be issued by states, cities, counties and various governmental agencies. As of 2010, over $3.7 trillion of municipal debt was outstanding and representing over three million bond issues. Municipal bonds are used to finance different types of projects ranging from hospitals and stadiums to general governmental operations. The majority of municipal bonds falls under two major categories, general obligation bonds and revenue bonds. General obligation bonds are typically secured by property tax revenues with either limited or unlimited pledges, where the bonds with unlimited pledges are typically viewed as the most secure type of bonds. The revenue bonds are issued to finance a particular project, such as a toll road, parking structure, etc. The principal and interest of these bonds are secured exclusively from the revenues of the particular project. Thus, these bonds typically do not have a claim on a municipality’s tax revenues.

Municipal bonds may be structured as a negotiated offering or as a competitive sale. The competitive bidding process is open to all underwriters and is awarded based on the overall interest cost. In negotiated sales, an underwriting syndicate is selected to purchase the bonds, which are then sold to investors. Most studies suggest that the competitive bidding process results in lower interest costs (Benson, 1979; Simonsen and Robbins, 1996).

Historically, the municipal bonds have had significantly lower rates of default than corporate bonds. Given a default rate of only one tenth of the corporate bond market, the municipal bonds are traditionally viewed as low credit risk investments. However, recent bankruptcies involving Detroit, Jefferson County and various other municipalities have increased both the frequency and monetary value of defaults to record levels[10]. The prevalence of higher defaults for municipal bonds that are unrated also belies the notion of municipal bonds as being safe investments.[11]

The municipal bond market is unique because of the prevalence of bond insurance. The bond insurers charge a premium and guarantee interest and principal payments in the event of default. Thus, insurance lowers the overall cost of borrowing for issuers and also alleviates the need for investors to monitor municipalities. Strong evidence is also provided in the literature that third-party insurance can decrease incentives for investors to monitor securities because insurers collect and analyze information on municipalities (Gore, Sachs, and Trzcinka, 2004). Bond insurance can be especially valuable because disclosures are less timely in the municipal bond market. Therefore, in the absence of insurance, investors may seek alternative sources of

[10] $9.02 billion in monetary defaults in 2014 compared with only $1.95 billion in 2012. http://www.bondbuyer.com/news/markets-buy-side/defaults-reached-record-in-2014-1069491-1.html

[11] While Moody's only notes 71 listed defaults from 1970 to 2011, there were 2,521 defaults during this same period for unrated municipal securities. http://libertystreeteconomics.newyorkfed.org/2012/08/the-untold-story-of-municipal-bond-defaults.html

information to evaluate the risk of municipal bonds. It is presented in this study that the social capital environment of municipalities is another source of information that is used to evaluate the risk of issuers. This study empirically tests whether the bonds issued by municipalities located in counties with high social capital have lower bond yields and yield spreads.

Social Capital

Social capital is generally perceived to be a by-product of social relations. It is considered to be a public good that benefits all members of a community (Putnam 1995; Woolcock 2001). Although many definitions are attached to the concept of social capital, it is generally defined in the form of social norms and social trust. Consistent with Jha and Chen (2014) and Jha (2013), this study adopts the definition by Woolcock (2001) that social capital reflects norms and networks that facilitate action and cooperation among individuals.

The basic premise of social capital is that individuals within a social network develop certain norms, which are associated with greater altruism, trustworthiness and the propensity to honor obligations. These norms become reinforced as individuals share common networks such as neighborhoods, churches, schools, and other civic associations. The norms developed in areas with high social capital encourage individuals in general to behave and make decisions consistent with these norms (North, 1990; Greif, 1994). It is argued that individuals deviating from social and ethical norms suffer self-guilt and isolation (e.g. Milgram, Bickman, and Berkowitz, 1969; Cialdini, Kallgren, and Reno, 1991). Therefore, it is emphasized in the literature that individuals from high social capital areas generally exhibit high social norms (e.g. Coleman, 1988; Spence et al., 2003; Melé, 2003).

There is evidence that high social capital can enhance firm performance (Fukuyama, 1995), improve leadership (Maak, 2007) and disclosure quality (Jha, 2013), and encourage ethical behavior (Pastoriza et al., 2008). Jha (2013) documents that firms headquartered in high social capital areas are associated with higher quality financial reports. Jha and Chen (2014) document that high social capital lowers risk and results in lower audit fees.

Hypothesis Development

Social Capital and Municipal Bonds

The borrowing cost associated with municipal bonds can arise from agency problems. In the municipal bond market, agency costs may occur because elected officials or public managers may be driven by self-interest and bondholders cannot generally observe their actions (Simonsen and Hill, 1998). The borrowing cost is expected to be high when investors perceive that municipal officials are corrupt, providing unreliable information or unwilling to honor obligations. This study expects that borrowing cost, represented by bond yields is also likely to be influenced by the level of social capital of the county in which the municipalities are located.

The borrowing cost of municipal bonds is likely to be lower in high social capital areas because officials from these counties are expected to be more honest (e.g. LaPorta et al., 1997). It has also been shown that governments from high social capital areas generally have efficient legal systems (Guiso et al., 2000). Putnam, Leonardi, and

Nanetti (1994) also document that Italian local governments in high social capital areas are associated with higher performance, whereas Fukuyama (1995) argues that high social capital enhances economic development. Lastly, Jha (2013) finds that social capital can influence the decision-making process in the corporate environment and shows that high social capital leads to better reporting quality.

Based on the above arguments, we expect that higher social capital environments that are associated with more efficient government, greater trustworthiness, and the propensity to honor obligations may reduce the borrowing cost of municipalities. In addition, a test was conducted to examine the impact of social capital on bond risk, which is represented by the yield spread between a municipal bond and similarly a matched Treasury bond. The following hypothesis was developed to test the expectations:

H1a: Ceteris paribus, the bonds issued by municipalities located in high social capital counties are associated with lower bond yields compared to the bonds issued by municipalities located in low social capital counties.

H1b: Ceteris paribus, the bonds issued by municipalities located in high social capital counties are associated with lower yield spread compared to the bonds issued by municipalities located in low social capital counties.

General Obligation Bonds versus Revenue Bonds

The majority of municipal bonds fall under two major categories, general obligation bonds and revenue bonds. General obligation bonds are secured by a pledge of property taxes, which local governments levy to meet the debt service requirements. Alm and Gomez (2008) claim that social capital plays an important role in determining an individual's intrinsic motivation to pay tax. Thus, in areas with high social capital, the issuers and their constituents might be more willing to honor their obligations through raising taxes.

Revenue bonds are issued to finance particular projects. The principal and interest of these bonds are secured exclusively from the revenues of the designated projects such as toll roads or stadium revenues. Since revenue bonds do not have a claim on the tax revenues of the issuing jurisdiction, social capital may have a limited impact on the cost of debt of revenue bonds. The following hypothesis was generated to differentiate the association of social capital with general obligations and revenue bonds:

H2: Ceteris paribus, the association between social capital and bond yields will be stronger for general obligation bonds compared to revenue bonds.

Bond Insurance and Social Capital

One of the benefits of bond insurance is that it alleviates the need for investors to monitor municipalities and their informational disclosures because this role is transferred to the insurers who guarantee principal and interest payments. High social capital environments that are associated with more efficient government, greater trustworthiness, and the property to honor obligations may reduce risk associated with municipal bonds. This also implies that high social capital reduces the need to monitor municipalities. This argument suggests that as a substitute, social capital may moderate the demand for insurance. This expectation was tested on the following hypothesis:

H3: Ceteris paribus, the bonds issued by municipalities located in high social capital counties are less likely to have insurance compared to the bonds issued by municipalities located in low social capital counties.

3.3. Data and Methodology

Sample Selection and Calculation of Variables

This study extracts bond information from the Mergent Municipal Bond Securities Database (Mergent) for the period from 1998 through 2012. The extracted information includes bond yields, offering size, bond offering date, bond insurance, the most recent bond rating (Moody, Standard and Poor's, and Fitch), bond type (general obligation, revenue, etc.), maturity date, optional call schedule, bank qualified indicator, and etc. Treasury information is extracted from the Center for Research in Security Prices (CRSP). Yield spread is calculated as the difference between the bond's yield to maturity and a benchmark Treasury yield using the daily CRSP fixed term indexes for the periods 1, 2, 5, 7, 10, 20, 30 years.

This study uses additional county-level demographic data from the United States Census. To control for macroeconomic level factors, the study collects data at the county level for income levels, consumer price index and etc. The Bond Buyer index was included to control for market risk[12]. The macroeconomic level data is matched with lag of one year from the bond issuance date. The main variables used in this study are defined in Appendix C.

This study uses county-level religiosity data from the Association of Religion Data Archive (ARDA). The religion adherents per capita are included to control for religiosity factors. To control for demographic fractionalization (Bergstresser et al. 2013), the study collects the religious fractionalization using the religion data from the

---

[12] The index consists of 20 general obligation bonds that mature in 20 years. The average rating of the 20 bonds is roughly equivalent to Moody's Aa2 rating.

ARDA. The religious fractionalization measure is generated on the basis of the Herfindahl index used in economics. Following Bergstresser et al. (2013), and Johnson and Grim (2013), a Religion Diversity Index (RDI) is created on the basis of eight religious groups: Mainline Protestant, Evangelical Protestant, Catholic, Eastern Orthodox, Other participating Christian groups (predominantly LDS (Mormon) congregations), Jewish, Muslim congregations, and unaffiliated. The RDI score is a version of the Herfindahl index and is inverted so that higher scores indicate higher diversity.

In the sample, there may be one or more rating agencies that rate a particular bond issue. Following Butler et al. (2009), the study uses Standard and Poor's rating if available. If Standard and Poor's rating is not available, Moody's rating is used. If both ratings are unavailable, Fitch's rating is used. Following Nanda and Singh (2004), the bond ratings are transformed into a numeric scale for regression analysis. The detailed classification scheme for the numerical score is provided in Appendix C.

Measure of Social Capital

Following Rupasingha, Goetz, and Freshwater (2006), this study constructs a social capital index for each county using principal component analysis. In the principal component analysis, four factors are considered: population, voter turnout, census response rates, and one aggregate variable. The aggregate variable is generated by grouping the number of county-level associations which include religious organizations, civic and social associations, business associations, political organizations, labor organizations, bowling centers, physical fitness facilities, public golf courses, sports clubs, charities and etc. The first principal component is extracted as the index of social capital. The social capital index is available for 3,108 counties for years 1997, 2005 and 2009. Following Jha and Chen (2014), the study linearly interpolates values to fill the missing years from 1998 to 2004 and 2006 to 2008. This study then linearly extrapolates values for years 2010 to 2012. The correlation between social capital levels in 1997 and 2009 is 88 percent. This is consistent with the idea that unlike physical and human capital, social capital is “sticky” (Anheier and Gerhards, 1995).

Table 3.3 presents the ten counties with the highest and lowest levels of social capital for the year 2005. Texas has seven counties, which are among either the top or bottom ten. Five counties have the lowest levels of social capital and two counties are among the highest.

Table 3.1: List of top ten low and high social capital counties

Rank	Low Social Capital Counties	High Social Capital Counties
1	Starr, TX	Edgefield, SC
2	Chattahoochee, GA	Loving, TX
3	Hidalgo, TX	Hooker, NE
4	Zapata, TX	Thomas, NE
5	Maverick, TX	Hinsdale, CO
6	Echols, GA	Nicollet, MN
7	Cameron, TX	Griggs, ND
8	La Paz, AZ	Motley, TX
9	Webb, TX	Lexington, VA
10	Yuma, AZ	Garfield, NE

Social capital index is merged with bond issuance data at the county level, lagged by one year. For instance, a bond issued in 2010 will be matched with county-level social capital information for 2009.

Descriptive Statistics

Table 3.1 provides descriptive statistics. There are 2,585,674 bond issues from 1998 through 2012, and 984,703 of them are merged with county-level social capital.

There are a total of 801,726 bond issues after removing missing value for variables including bond yield, bond size, issue size, maturity, GO bond, callable provision, bank qualified. There are 351,677 of 801,726 bonds with ratings from the major rating agencies. The average bond rating is roughly 5.03, which is equivalent to an S&P rating between AA and AA-. The average yield to maturity at the time of the bond issue is 3.7%. The average yield spread at the time of the bond issue is -0.087 percent. A Pearson correlation matrix is provided in Table 3.2. The correlation coefficients are consistent with the main predictions. Social capital is negatively associated with bond yields, yield spreads, and the usage of bond insurance. Consistent with expectations, social capital is positively correlated with the income levels.

Table 3.2: Summary Statistics

Variable	Full Sample							
	N	Mean	Std. Dev.	Min	Max	1st Quartile	2nd Quartile	3rd Quartile
Bond Yield	801726	3.68	1.3	0	100	2.96	3.88	4.5
Yield Spread	767013	-0.09	0.9	-5.39	96.42	-0.7	-0.24	0.39
Social Capital	801726	-0.25	1.09	-4.02	19.14	-1	-0.3	0.32
Volume-weighted Price	1458331	100.65	9.61	0.01	130	99.94	101.75	104.9
Bond Buyer Index	801347	4.66	0.52	3.27	6.09	4.33	4.62	5.05
Log(Bond size)	801721	13.11	1.57	0.96	22.9	12.04	13.01	14.09
Log(Issue size)	801726	16.03	1.53	1.39	21.88	15.01	15.94	17.03
Maturity	801726	9.44	6.25	0.02	55.37	4.5	8.43	13.3
Rating	801726	1.65	3.05	-1	6	-1	-1	5
Nonrated dummy	801726	0.56	0.5	0	1	0	1	1
Insurance	801726	0.44	0.5	0	1	0	0	1
GO Bond	801726	0.6	0.49	0	1	0	1	1
Competitive Bid	801726	0.32	0.47	0	1	0	0	1
Callable Provision	801726	0.43	0.5	0	1	0	0	1
Bank Qualified	801726	0.46	0.5	0	1	0	0	1
Log(income)	801726	9.75	0.31	8.79	10.99	9.56	9.73	9.87
Religion Adherents per Capita	801726	0.52	0.14	0.02	1.2	0.42	0.52	0.6
Religion Diversity Index	801726	6.99	1.09	0.3	8.7	6.5	7.2	7.7

This table reports descriptive statistics for key variables. See Appendix C for variable definitions. N is the number of observations. Mean is the average value, min is the minimum and max is the maximum value. Bond yield and maturity are restricted to positive values.

Table 3.3: Correlation

Variable	Bond Yield	Yield Spread	Social Capital	Bond Buyer Index	Log Bond size	Log Issue size	Maturity	Rating	Nonrated dummy	Insurance	GO Bond	Comp -etitive	Callable Provision	Bank Qualified	Log Income	Religion Adherents per Capita	Religion Diversity Index
Bond Yield	1																
Yield Spread	0	1															
Social Capital	-0.02	-0.05	1														
Bond Buyer Index	0.53	-0.3	0.05	1													
Log Bond size	0.11	0.06	-0.15	-0.06	1												
Log Issue size	0.06	-0.04	-0.2	-0.03	0.77	1											
Maturity	0.61	0.11	-0.08	0.06	0.35	0.19	1										
Rating	0.05	0.19	-0.1	-0.29	0.34	0.28	0.48	1									
Nonrated dummy	-0.06	-0.22	0.11	0.3	-0.33	-0.27	-0.49	-0.98	1								
Insurance	0.14	-0.31	-0.14	0.16	0.07	0.15	0.07	0	-0.01	1							
GO Bond	-0.12	-0.09	0.05	-0.02	-0.17	-0.23	-0.07	0.01	0.01	0.01	1						
Competitive	-0.23	0.16	0.01	-0.31	-0.02	-0.04	-0.02	0.18	-0.18	-0.17	0.13	1					
Callable Provision	0.48	0.09	-0.01	0.03	0.17	0.06	0.74	0.39	-0.4	0.03	-0.01	0.01	1				
Bank Qualified	-0.13	-0.04	0.15	-0.02	-0.51	-0.63	-0.12	-0.17	0.15	-0.08	0.27	0.06	0	1			
Log Income	-0.03	0.09	0.18	-0.08	0.27	0.36	0.04	0.11	-0.11	-0.09	-0.08	0.07	0.01	-0.26	1		
Religion Adherents per Capita	-0.02	0.02	0.2	0	0	-0.02	-0.02	-0.02	0.02	-0.12	-0.04	0.08	0.01	0.03	0.17	1	
Religion Diversity Index	0	0.05	0.39	0	0	0	-0.01	-0.03	0.03	-0.09	0.03	0.05	0.03	0.01	0.41	0.33	1

3.4. Research Methodology

Social Capital and Borrowing Costs

The study uses the ordinary least squares (OLS) regression analysis to test the hypotheses and uses the bond yield, reflecting cost of municipal bonds, as a dependent variable. The variable of interest in this specification is the county-level social capital. The coefficient on social capital is expected to be negative and significant, indicating that high social capital is associated with lower bond yields, reflecting low bond costs.

$$Bond\ Yield = \beta_0 + \beta_1 Social\ Capital + \beta_2 Bond\ Buyer\ Index + \beta_3 Log\ Bond\ Size + \beta_4 Log\ Issue\ Size + \beta_5 Maturity + \beta_6 Rating + \beta_7 Nonrated + \beta_8 Insurance + \beta_9 General\ Obligation\ Bond + \beta_{10} Competitive\ Bid + \beta_{11}\ Callable + \beta_{12} Bank\ Qualified + \beta_{13}\ Log\ Income + \beta_{14} Religion\ Adherents\ per\ Capita + \beta_{15} Religion\ Diversity\ Index + Year\ Indicators + State\ Indicators + \varepsilon \qquad (3.1)$$

The study uses the bond-level control variables that have been used by Hastie (1972), Blackwell and Kidwell (1988), Kao and Wu (1994), and Nanda and Singh (2004). These include controls for issue size, maturity, rating, the method of sale, the type of bond, callable bonds, and bank qualified bonds. The issue size is measured by taking the natural log of the principal amount of the bond's original offering. Maturity is the life of a bond in years and calculated as the days difference between the maturity date and the offering date divided by 365. Bond ratings are the alphanumeric conversions of ratings issued by the rating agencies and used to control for credit risk. The bonds that do not have the rating information take a value of negative 1. The offering type indicates the method of sale and takes a value of 1 if the issue sale is competitive and 0 value if the sale is negotiated. Insurers provide a guaranty of principal and interest payments for municipal bonds in the event of default. The credit rating on bonds with insurance is the insurers' creditworthiness instead of the municipality's underlying creditworthiness. The study includes an indicator variable, insurance, which takes a value of 1 if the bond is insured and 0 if the bond is not insured. Both state and year fixed effects are included. The standard errors of the regression are clustered at the county level.

The county-level demographic controls include three variables: income per capita, religion adherents per capita, and the religion diversity index. Income per capita is deflated by the Consumer Price Index. Recent finance and accounting research have indicated that religiosity may also influence the cost of corporate debt (Jiang, Li, and Qian 2013). Although religiosity and social capital are different constructs, religion is a component in social capital (e.g. Jha and Chen 2014). In order to examine whether the findings in this study are incremental to the effect of religiosity, this study controls for the religious adherents per capita. Lastly, Bergstresser et al. (2013) find that bonds issued from more ethnically and religiously fractionalized counties have higher yields. To demonstrate that the results in this study are incremental to their findings, this study controls for fractionalization[16].

A comparable analysis was conducted to examine the impact of social capital on the borrowing costs before and after the downgrade of major municipal bond insurers. It is expected that the influence of social capital on the borrowing costs to be stronger when the bond insurers' credit worthiness deteriorate. The downgrade of the major municipal bond insurers was precipitated by their exposure to subprime mortgages. This increased exposure to risk led to the loss of triple A ratings for MBIA and AMBAC. The loss of triple A ratings began in June 2008 when Moody's downgraded AMBAC's credit rating three notches to Aa3 from Aaa. The credit

[16] The test includes religious fractionalization. Ethnic fractionalization is not included in the main model since it is highly correlated with social capital. Their correlation is 0.54.

rating of AMBAC continued to decline afterwards and the insurer eventually defaulted. The MBIA also lost its triple A rating in 2008 and it was eventually given a speculative grade rating.

The study argues that this deterioration in the ratings of the bond insurers motivated investors to explore alternative sources to evaluate the risk of municipal bonds. Without high quality insurance that serves to monitor municipalities, investors may rely more on social capital as a source of information, and consequently they would be willing to accept lower bond yields. The study refers to the period from June 19th, 2007 to June 18th, 2008 as the pre-downgrade period, and the period from June 19th, 2008 to June 19th, 2009 as the post-downgrade period.

Social Capital and Bond Risk

The study examines the yield spread to capture the risk component in the pricing of municipal bonds using the equation 3.2. The study uses yield spread as dependent variable in the OLS regression analysis and the control variables in this equation are same as in equation 3.1. The coefficient on social capital is expected to be negative and significant, indicating that high social capital is associated with lower yield spreads, reflecting low risk.

$$\begin{aligned}&Yield\ Spread = \beta_0 + \beta_1 Social\ Capital + \beta_2 Bond\ Buyer\ Index +\\&\beta_3 Log\ Bond\ Size + \beta_4 Log\ Issue\ Size + \beta_5 Maturity + \beta_6 Rating +\\&\beta_7 Nonrated + \beta_8 Insurance + \beta_9 General\ Obligation\ Bond +\\&\beta_{10} Competitive\ Bid + \beta_{11}\ Callable + \beta_{12} Bank\ Qualified + \beta_{13}\ Log\ Income +\\&\beta_{14} Religion\ Adherents\ per\ Capita + \beta_{15} Religion\ Diversity\ Index +\\&Year\ Indicators + State\ Indicators +\ \varepsilon \qquad (3.2)\end{aligned}$$

The study also conducts a test to compare the association between social capital and yield spread for the pre-downgrade period with the post-downgrade period. The association is expected to be stronger for the post-downgrade period compared to the pre-downgrade period.

Bond Insurance and Social Capital

It is anticipated that demand for insurance decreases if municipalities are perceived to be trustworthy and likely to honor obligations. It is expected that the municipalities from counties with high social capital are less likely to issue bonds with insurance. To test this, this study utilizes a logistic regression in which the dependent variable is an indicator variable taking a value of 1 if the bond is insured and 0 otherwise. The coefficient on the variable of social capital is expected to be negative and significant, indicating that high social capital regions are less likely to have bonds that are insured.

$$\begin{aligned}&Insurance = \beta_0 + \beta_1 Social\ Capital + \beta_2 Log\ Issue\ Size + \beta_3 Maturity + \\ &\beta_4 General\ Obligation\ Bond + \beta_5 Competitive\ Bid + \beta_6 Rating + \\ &\beta_7 Nonrated + \beta_8 Callable + \beta_9 Bank\ Qualified + \beta_{10} Log\ Income + \\ &\beta_{11} Religion\ Adherents\ per\ Capita + \beta_{12} Religion\ Diversity\ Index + \\ &Year\ Indicators + State\ Indicators + \varepsilon \qquad (3.3)\end{aligned}$$

This test controls for issuer size, maturity, rating levels, offering type, and county-level demographic: income per capita, religion adherents per capita, and religious diversity. Similar to equation 3.1, state and year fixed effects are included. Standard errors are clustered at the county-level.

3.5. Results

Association between Social Capital and Borrowing Cost

The results of tests on the association between social capital and borrowing cost are presented in Table 3.4. The coefficient for social capital (-0.22) in Model 1 is negative and significant at the 5 percent level. These results indicate that the cost of borrowing, reflected by the bond yield, is lower for municipalities located in counties with high social capital. The results specifically show that an issuer in a county with social capital in the 75th percentile pays 3 percent less in bond yields compared to an issuer in a county with social capital in the 25th percentile[17], ceteris paribus. The results in Model 1 in Table 3.4 show that the coefficient (-0.133) of credit rating is negative and significant, indicating that default risk is priced into yields.

The results of Model 1 also indicate that the coefficients of control variables are generally consistent with expectations. The bond yields are higher for issues with higher risk characteristics; these include bonds with longer maturities, low ratings, and without bond insurance. Consistent with previous research, the results indicate that the negotiated offerings have higher interest costs (Benson, 1979; Simonsen and Robbins, 1996). Bonds with callable provisions and without bank qualification may also have higher yields. The general obligation bonds also have lower yields since the taxing authority of the issuing government backs them.

[17] 1-(exp (-0.022*0.3291)/exp(-0.022*-0.9998))=2.9%

Table 3.4: Impact of Social Capital on Bond Yields

Variable	Predicted Sign	Model 1 Full Sample		Model 2 Pre-Period		Model 3 Post-Period	
		Coefficient	T Stat	Coefficient	T Stat	Coefficient	T Stat
Social Capital	**-**	**-0.022*** **	**-2.72**	**-0.028** **	**-2.32**	**-0.055*** **	**-2.65**
Bond Buyer Index	+	0.730***	55.29	0.009	0.22	0.800***	19.78
Log(Bond Size)	+/-	-0.064***	-9.38	-0.044***	-5.26	-0.113***	-9.3
Log(Issue Size)	+/-	-0.043***	-7.45	-0.071***	-6.96	-0.067***	-4.86
Maturity	+	0.112***	43	0.074***	39.4	0.137***	32.98
Rating	-	-0.133***	-21.23	-0.079***	-8.37	-0.237***	-11.32
Nonrated Dummy	+/-	-0.776***	-16.62	-0.424***	-6.71	-1.693***	-12.13
Insurance	-	-0.144***	-14.68	-0.082***	-4.53	-0.099***	-3.03
GO Bond	-	-0.152***	-10.03	-0.191***	-10.03	-0.210***	-6.27
Competitive Bid	-	-0.104***	-6.74	-0.078***	-3.81	-0.112***	-2.59
Callable	+	0.182***	11.63	0.169***	10.52	0.241***	8.57
Bank Qualified	-	-0.321***	-16.57	-0.409***	-15	-0.421***	-9.86
Log(Income)	-	0.008	0.28	0.041	1	0.052	0.67
Religion Adherence per Capita	-	-0.106***	-2.43	-0.079	-1.17	-0.209	-1.59
Religion Diversity Index	+	0.009	1.48	0.007	0.66	0.047***	2.45
Intercept		0.752	2.31	5.239	11.11	2.1	2.4
Year, State Dummies		Included		Included		Included	
N		801,342		50,194		45,266	
R-squared		73.69%		55.68%		71.04%	

This table shows the regression results for models 1-3. Model 1 is the main model to examine the impact of social capital on borrowing cost. Model 2 examines the impact of social capital on borrowing cost one year before rating downgrade of insurers. Model 3 examines the impact of social capital one year after rating downgrade of insurers. All models include state and year fixed effects. All regressions are clustered by county. See Appendix C for variable descriptions. N is the sample size of the regression. R-squared represents a goodness of fit measure. Bond yields and maturity are restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

Additionally, the study conducts a test on the pre- and post-downgrade period for bond insurers. The impact of social capital on bond yields is expected to be stronger for the post-downgrade period. The results in Models 2 and 3 of Table 3.4 show the association between bond yield and social capital values for the pre- and post-period, respectively. The results show that in the post-downgrade period from June 2008 to June 2009, the coefficient of social capital on bond yields is -0.055, whereas it is -0.028 in the pre-downgrade period. The difference in the coefficients for two periods is significant as indicated by an F-test. In terms of economic values, in the post-downgrade period, municipalities in a county with social capital in the 75th percentile have paid 7 percent less in bond yields than municipalities located in a county with social capital in the 25th percentile.[18] In the pre-downgrade period, an issuer in a county with social capital in the 75th percentile paid 3.7 percent less in bond yields than an issuer in a county with social capital in the 25th percentile.[19] These results thus indicate that social capital played a more important role in evaluation after the major insurers were downgraded.[20] Thus, social capital acted as a substitute for bond insurance. These findings provide addition support to the hypothesis H1a.

Association between Social Capital and Bond Risk

Further, the study examines the association between social capital and the yield spread to better capture the risk component of municipal bonds. Table 3.5 presents the results of tests on their association. The variable of credit rating is

---

[18] 1-(exp (-0.055*0.3291)/exp(-0.055*-0.9998))=7.0%

[19] 1-(exp (-0.028*0.3291)/exp(-0.028*-0.9998))=3.7%

[20] The results remain unchanged when the pre-downgrade period and the post-downgrade period are extended to two, three, and four years. In separate analysis, the study also examines the interaction between social capital and the post downgrade period and finds consistent results.

negative and significant in the tests, indicating that the default risk significantly explains yields in the municipal bond market. The results in Model 1 show that the coefficient (0.023) for social capital is negative and significant at the 5-percentile level, which indicates that the bond risk is lower when the municipalities are located in counties with high social capital. These results support the hypothesis H1b.

Table 3.5: Impact of Social Capital on Yield Spread

Variable	Predicted Sign	Model 1 Full Sample		Model 2 Pre-Period		Model 3 Post-Period	
		Coefficient	T Stat	Coefficient	T Stat	Coefficient	T Stat
Social Capital	**-**	**-0.023*** **	**-2.87**	**-0.032*** **	**-2.74**	**-0.054*** **	**-2.74**
Bond Buyer Index	+	0.271***	22.83	0.125**	2.09	1.745***	36.06
Log(Bond Size)	+/-	0.019***	3.93	0.061***	7.14	0.001	0.08
Log(Issue Size)	+/-	-0.109***	-18.28	-0.129***	-10.48	-0.086***	-5.93
Maturity	+	0.030***	37.32	0.007***	5.52	0.022***	8.37
Rating	-	-0.143***	-17.86	-0.086***	-7.63	-0.286***	-12.48
Nonrated Dummy	+/-	-0.639***	-12.92	-0.214***	-3.15	-1.281***	-9.13
Insurance	-	-0.151***	-17.3	-0.202***	-10.76	0.027	0.85
GO Bond	-	-0.184***	-12.02	-0.126***	-5.02	-0.240***	-6.37
Competitive Bid	-	-0.126***	-6.78	-0.110***	-4.7	-0.085**	-2.16
Callable	+	0.050***	6.29	-0.052***	-3.1	0.045*	1.76
Bank Qualified	-	-0.322***	-15.78	-0.332***	-10.29	-0.300***	-5.36
Log(Income)	-	0.036	1.16	0.096**	2.1	0.165**	2.17
Religion Adherence per Capita	-	-0.114***	-2.43	-0.13	-1.6	-0.194	-1.59
Religion Diversity Index	+	0.01	1.4	0.021	1.65	0.045**	2.35
Intercept		1.71	5.11	0.531	1.04	-6.683	-7.88
Year, State Dummies		Included		Included		Included	
N		766,653		48,012		43,050	
R-squared		54.16%		24.11%		47.18%	

This table shows the regression results for models 1-3. Model 1 is to examine the impact of social capital on yield spread on the full sample. Model 2 examines the impact of social capital on yield spread one year before rating downgrade of insurers. Model 3 examines the impact of social capital one year after rating downgrade of insurers. All models include state and year fixed effects. All regressions are clustered by county. See Appendix C for variable descriptions. N is the sample size of the regression. R-squared represents a goodness of fit measure. Bond yields and maturity are restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

The study also presents the results on bond yield for pre- and post-downgrade periods of bond insurers in Models 2 and 3 respectively. The impact of social capital on the bond risk is expected to be stronger during the post-downgrade period compared to the pre-downgrade period. The results show that in the post-downgrade period the coefficient of social capital on yield spread is -0.054 (Model 3), while it is -0.032 in the pre-downgrade period (Model 2). In terms of economic values, in the post-downgrade period, the municipalities in a county with social capital in the 75th percentile have 5.7 percent less in risk premium than municipalities located in a county with social capital in the 25th percentile.[21] In the pre-downgrade period, the municipalities in a county with social capital in the 75th percentile have 3.4 percent less in risk premium than an issuer in a county with social capital in the 25th percentile.[22] The results of F-test show that difference in the coefficients for two periods is significant. These results thus indicate that the impact of social capital on yield spread became stronger after insurers were downgraded. The findings additionally provide support for hypothesis H1b.

Impact of Social Capital on General Obligation Bonds Versus Revenue Bonds.

Separate tests were conducted to examine the impact of social capital on general obligation bonds and revenue bonds. In the sample, 59.88% of bonds are the general obligation bonds. The results of Model 1 of Table 3.6 show that the association between social capital and bond yields for the general obligation bonds is negative and statistically significant at the 3-percentile level, which is greater than the overall sample. Similar to the main findings, Models 2 and 3 of Table 3.6 show that the economic impact of social capital is greater in the post period when insurers lost

---

[21] 1-(exp (-0.054*0.3887)/exp(-0.054*-0.6986))=5.7%

[22] 1-(exp (-0.032*0.3887)/exp(-0.032*-0.6986))=3.4%

their triple A ratings. The results support the hypothesis H2 and indicate that municipalities located in a higher social capital county issue general obligation bonds with lower bond yields. The results of Table 3.7 present that the association between social capital and yield spreads for the general obligation bonds is significant. Models 2 and 3 of Table 3.7 indicate the impact of social capital becomes stronger in the post-downgrade period. The results support the hypothesis H2.

Table 3.6: Impact of Social Capital on Borrowing Cost for General Obligation Bonds

Variable	Predicted Sign	Model 1 Full Sample		Model 2 Pre-Period		Model 3 Post-Period	
		Coefficient	T Stat	Coefficient	T Stat	Coefficient	T Stat
Social Capital	**-**	**-0.027*****	**-3.51**	**-0.032*****	**-2.56**	**-0.077*****	**-3.08**
Bond Buyer Index	+	0.713***	56.69	0.013	0.28	0.817***	16.87
Log(Bond Size)	+/-	-0.067***	-9.17	-0.063***	-6.48	-0.120***	-7.34
Log(Issue Size)	+/-	-0.047***	-5.28	-0.067***	-4.77	-0.078***	-4.17
Maturity	+	0.129***	70.52	0.077***	33.6	0.150***	38.68
Rating	-	-0.100***	-10.33	-0.059***	-5.6	-0.192***	-6.63
Nonrated Dummy	+/-	-0.569***	-8.46	-0.315***	-4.37	-1.405***	-7.78
Insurance	-	-0.079***	-6.32	-0.023	-0.94	-0.113**	-2.13
Competitive Bid	-	-0.042***	-3.72	-0.011	-0.43	-0.037	-0.75
Callable	+	0.068***	3.31	0.130***	6.52	0.154***	4.57
Bank Qualified	-	-0.303***	-11.27	-0.382***	-10.65	-0.419***	-6.91
Log(Income)	-	0.099**	1.98	0.061	0.82	0.037	0.27
Religion Adherence per Capita	-	-0.147***	-3.19	-0.119	-1.53	-0.282*	-1.9
Religion Diversity Index	+	0.011*	1.73	0.003	0.29	0.057***	2.54
Intercept		-0.623	-1.08	4.967	6.21	1.804	1.19
Year, State Dummies		Included		Included		Included	
N		480,248		31,120		29,536	
R-squared		78.04%		57.95%		73.74%	

This table shows the OLS regression results of the sample general obligation bonds based on equation 3.1. Model 1 is to examine the impact of social capital on borrowing cost for general obligation bonds for the period 1998 to 2012. Model 2 examines the impact of social capital on borrowing cost of general obligation bonds one year before rating downgrade of insurers. Model 3 examines the impact of social capital for general obligation bonds one year after rating downgrade of insurers. All models include state and year fixed effects. All regressions are clustered by county. See Appendix C for variable descriptions. N is the sample size of the regression. R-squared represents a goodness of fit measure. Bond yields and maturity are restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

Table 3.7: Impact of Social Capital on Yield Spread for General Obligation Bonds

Variable	Predicted Sign	Model 1 Full Sample Coefficient	Model 1 Full Sample T Stat	Model 2 Pre-Period Coefficient	Model 2 Pre-Period T Stat	Model 3 Post-Period Coefficient	Model 3 Post-Period T Stat
Social Capital	**-**	**-0.028*** **	**-3.44**	**-0.034*** **	**-2.52**	**-0.069*** **	**-2.87**
Bond Buyer Index	+	0.287***	18.31	0.02	0.34	1.751*	33.21
Log(Bond Size)	+/-	-0.006	-1.05	0.020**	2.04	-0.017	-1.05
Log(Issue Size)	+/-	-0.094***	-10.7	-0.110***	-8.28	-0.082***	-4.12
Maturity	+	0.030***	44.14	0.006***	3.96	0.024***	7.35
Rating	-	-0.107***	-10.9	-0.058***	-4.55	-0.251***	-7.23
Nonrated Dummy	+/-	-0.476***	-8.01	-0.065	-0.79	-1.100***	-5.21
Insurance	-	-0.077***	-6.05	-0.090***	-3.95	0.034	0.74
Competitive Bid	-	-0.066***	-4.35	-0.059**	-2.28	-0.015	-0.32
Callable	+	0.014*	1.76	-0.093***	-4.89	-0.011	-0.39
Bank Qualified	-	-0.292***	-10.91	-0.315***	-8.38	-0.295***	-3.95
Log(Income)	-	0.136***	2.38	0.167***	2.37	0.146	1.08
Religion Adherence per Capita	-	-0.146***	-2.85	-0.152*	-1.71	-0.269*	-1.93
Religion Diversity Index	+	0.004	0.61	0.015	1.04	0.055***	2.63
Intercept		0.154	0.23	0.352	0.44	-6.651	-4.53
Year, State Dummies		Included		Included		Included	
N		453,927		29,403		27,791	
R-squared		57.78%		24.01%		49.60%	

This table shows the regression results for the sample of general obligation bond based on equation 3.2. Model 1 is to examine the impact of social capital on yield spread for general obligation bonds for the period 1998 to 2012. Model 2 examines the impact of social capital on yield spread one year before rating downgrade of insurers. Model 3 examines its impact one year after rating downgrade of insurers. All models include state and year fixed effects. All regressions are clustered by county. See Appendix C for variable descriptions. N is the sample size of the regression. R-squared represents a goodness of fit measure. Bond yields and maturity are restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

On the other hand, the results do not show that social capital influences the cost of borrowing for revenue bonds (Table 3.8, Table 3.9). Since the revenue bonds are secured specifically by the income generated from projects, the social capital of a county in which the municipalities are located will have less effect on these bonds. These results support hypothesis H2

Table 3.8: Impact of Social Capital on Borrowing Cost for Revenue Bonds

		Full Sample	
Variable	Predicted Sign	Coefficient	T Stat
Social Capital	**-**	**0.003**	**0.24**
Bond Buyer Index	+	0.751***	22.85
Log(Bond Size)	+/-	-0.026***	-2.54
Log(Issue Size)	+/-	-0.056***	-5.06
Maturity	+	0.098***	43.38
Rating	-	-0.137***	-10.11
Nonrated Dummy	+/-	-0.802***	-9.24
Insurance	-	-0.214***	-8.28
Competitive Bid	-	-0.126***	-4.33
Callable	+	0.259***	11.37
Bank Qualified	-	-0.300***	-13.84
Log(Income)	-	-0.139***	-3.51
Religion Adherence per Capita	-	-0.002	-0.02
Religion Diversity Index	+	0.011	1.12
Intercept		2.157	4.94
Year, State Dummies		Included	
N		134,628	
R-squared		69.86%	

This table shows the regression results for the sample of revenue bonds based on equation 3.1. The model examines the impact of social capital on borrowing cost for revenue bonds for the period 1998 to 2012. The model includes state and year fixed effects. All regressions are clustered by county. See Appendix C for variable descriptions. N is the sample size of the regression. R-squared represents a goodness of fit measure. Bond yields and maturity are restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

Table 3.9: Impact of Social Capital on Yield Spread for Revenue Bonds

		Full Sample	
Variable	Predicted Sign	Coefficient	T Stat
Social Capital	**-**	**0.001**	**0.05**
Bond Buyer Index	+	0.254***	8.12
Log(Bond Size)	+/-	0.064***	5.45
Log(Issue Size)	+/-	-0.122***	-10.38
Maturity	+	0.026***	13.09
Rating	-	-0.142***	-11.45
Nonrated Dummy	+/-	-0.640***	-7.75
Insurance	-	-0.201***	-8.14
Competitive Bid	-	-0.137***	-3.76
Callable	+	0.070***	4.65
Bank Qualified	-	-0.300***	-11.85
Log(Income)	-	-0.112***	-2.66
Religion Adherence per Capita	-	0.007	0.09
Religion Diversity Index	+	0.016	1.59
Intercept		2.964	6.72
Year, State Dummies		Included	
N		130,764	
R-squared		49.12%	

This table shows the regression results of the sample of revenue bonds for the period 1998 to 2012 based on equation 3.2. The model examines the impact of social capital on yield spread for revenue bonds. The model include state and year fixed effects. All regressions are clustered by county. See Appendix C for variable descriptions. N is the sample size of the regression. R-squared represents a goodness of fit measure. Bond yields and maturity are restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

The Effect of Social Capital on the Demand for Bond Insurance

The study examines whether social capital influences the demand for bond insurance. Model 1 of Table 3.10 shows that the coefficient of social capital is -0.051, which is statistical significant at the 10% level. This result indicates that the likelihood of bonds carrying insurance is lower in municipalities located in high social capital counties. A one standard deviation change in social capital is associated with a decrease in the likelihood of insurance by 1.15 times.[23]

Table 3.10: Logistic Regression on the Likelihood of Insurance

Dependent Variable= Bond Insurance			
Variable	Predicted Sign	Coefficient	Z Stat
Social Capital	**-**	**-0.097***	**-1.84**
log(Issue Size)	+	0.395***	5.35
Maturity	+/-	-0.022***	-7.14
Rating	-	-0.333***	-6.42
Nonrated Dummy	+/-	-2.964***	-9.62
GO Bond	+/-	-0.003	-0.03
Competitive	+/-	0.058	0.81
Callable	-	-0.197***	-5.17
Bank Qualified	+/-	0.420***	3.63
Log(Income)	-	-1.351***	-7.23
Religion Adherence per Capita	-	-0.677***	-2.46
Religion Diversity Index	+/-	0.042	0.99
Intercept		8.109	3.46
Year, State dummies		Included	
N		581,399	
Pseudo R-squared		18.82%	

This table shows the logistic regression results for the impact of social capital on the demand of insurance. Models include state and year fixed effects. The regression is clustered by county. See Appendix C for variable descriptions. N is the sample size of the regression. Pseudo R-squared represents a goodness of fit measure. Bond yields and maturity are restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

The results suggest that high social capital moderates the demand for insurance because municipals bonds from high social capital areas are perceived to have lower risk. These findings provide evidence in support of hypothesis H3.

[23] Exp(0.1366)=1.15

3.6. Robustness Tests

Effect of Social Capital on Trade Price in Secondary Market

It is possible that the decision to issue new debt may confound the results. To overcome this weakness, an additional test was conducted on the same bond issues by comparing the price of the same bond issues in the secondary market for the pre- and post-downgrade periods. This analysis has the advantage of identifying how social capital influences the risk perceptions for a constant sample of bonds. It is expected that bonds issued by the municipalities in counties with high social capital will have higher prices in the secondary market. The positive association between bond prices and social capital is expected to be stronger in the post-downgrade period compared to the pre-downgrade period.

This study utilizes the OLS regression and uses the volume-weighted trade price as the dependent variable. The coefficient on the variable of social capital is expected to be positive and significant. This test is conducted on bonds that have been traded during both the pre-downgrade and the post-downgrade periods.

$$\begin{aligned} Price &= \beta_0 + \beta_1 Social\ Capital + \beta_2 Log\ Bond\ Size + \beta_3 Maturity + \\ &\beta_4 General\ Obligation\ Bond + \beta_5 Competitive\ Bid + \beta_6 Rating + \\ &\beta_7 Nonrated + \beta_8 Callable + \beta_9 Bank\ Qualified + \beta_{10} Sinking\ Fund + \\ &\beta_{11} Log\ Income + \beta_{12} Religion\ Adherents\ per\ Capita + \\ &\beta_{13} Religion\ Diversity\ Index + \varepsilon \end{aligned} \tag{3.4}$$

Following previous studies (Harris and Piwowar, 2006; Schultz, 2012), the test controls for credit quality, bonds with callable options, credit enhancement, bond size, maturity, offering type, and bonds with sinking funds. Additionally, the test controls for county-level demographic variables: income per capita, religious adherents per capita, and religious diversity. Standard errors are clustered at the bond-level.

This study obtains municipal bond trade data during the period of June 2007 and June 2009 from the Municipal Securities Rulemaking Board (MSRB) Historical Transaction Database. The database provides the price of the trade, the CUSIP number of the issue traded, security description, coupon, trade date, maturity date, an indicator showing whether the trade was initiated as a purchase from a customer, a sale from a customer, or an interdealer transaction. Following Bessembinder et al. (2009), this study calculates a volume-weighted trade price for each bond on each date. The trades are restricted to only customer imitated buy orders (Downing and Zhang, 2004). The volume-weighted trade price matches with its bond characteristics from the Mergent dataset by CUSIP number. The study examines the same bonds in the pre-downgrade period and the post-downgrade period. The final sample consists of 1,458,331 total trades. The average bond price is 101.46 during pre-downgrade period, and 99.90 during post-downgrade period.

The results are contained in Table 3.11. The results, based on the 694,822 trades in the pre-downgrade and 763,509 trades in the post-downgrade period, show that there is a positive and statistically significant association between trade price and social capital values. Additionally, the results show that the coefficient of social capital is lower in the pre-downgrade period when compared to the post-downgrade period (0.269 vs. 0.475). The difference in the coefficients is significant as indicated by the F-test. Holding virtually all other bond related factors constant, the result shows that the trade price of municipal bonds is more associated with social capital values in the post-downgrade period. This result is consistent with the main analysis.

Table 3.11: The Impact of Social Capital on Trade Price of the Secondary Market

	Pre-downgrade Period		Post-downgrade Period	
Variable	Coefficient	T Stat	Coefficient	T Stat
Social Capital	**0.269*** **	**3.1**	**0.475*** **	**4.09**
Bond Buyer Index	1.510***	7.04	3.064***	9.51
Log(Issue Size)	1.245***	14.91	1.329***	13.38
Maturity	-0.533***	-12.69	-0.795***	-16.22
Rating	0.11	1.03	1.692***	11.5
Nonrated Dummy	0.783	1.24	9.760***	11.04
Insurance	-0.276**	-2.13	-0.14	-0.61
GO Bond	-0.689***	-3.79	0.391*	1.74
Competitive Bid	0.788***	5.58	0.724***	3.97
Callable	5.475***	9.98	4.927***	6.97
Bank Qualified	0.953***	4.62	0.419	1.48
Sinking Fund	4.206***	9.02	4.344***	8.24
Log(Income)	0.095	0.29	-0.426	-1
Religion Adherence per Capita	-0.705	-0.76	0.268	0.29
Religion Diversity Index	-0.165	-1.06	-0.270*	-1.88
Intercept	76.293	36.4	67.108	17.15
Year, State Dummies	Included		Included	
N	694,822		763,509	
R-squared	12.04%		18.19%	

This table shows the regression results of the impact of social capital on volume-weighted trade price on the secondary market. Pre-period model examines the impact of social capital one year before rating downgrade of insurers. Post-period model examines the impact of social capital one year after rating downgrade of insurers. All regressions are clustered by issuance. See Appendix C for variable descriptions. N is the sample size of the regression. R-squared represents a goodness of fit measure. Maturity is restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

Joint Effect of Social Capital and Downgrades on the Bond Price

The credit rating downgrades of the major bond insurers may motivate investors to utilize social capital as an alternative source of information for risk evaluation. This study conducts an additional test by examining the impact of social capital on the secondary market price of insured bonds after the downgrade of the credit worthiness of insurers. The study evaluates the joint effect of social capital and downgrades on the traded price in the secondary market for insured bonds versus non-insured bonds using a three-way interaction variable between social capital, insurance, and the post-downgrade period. The post-period is coded as 1 for the period after downgrade of the insurers and 0 otherwise. The variable of insurance is coded as 1 if the bond is insured and 0 otherwise. This study includes this three-way interaction along with all necessary two-way interactions in equation 3.4.

The regression results are presented in Table 3.12. The results based on 632,438 trades for general obligation bonds show that both social capital and the three-way interaction variable are positive and statistically significant, indicating that for insured bonds, social capital had a greater affect on trade prices in the post period. This result indicates that social capital played an important role in the pricing of the insured bonds as the major bond insurers deteriorated in credit quality. These results support the main analyses that social capital may be a substitute for bond insurance.

Table 3.12: The Impact of Social Capital on Trade Price of the Secondary Market for General Obligation Bonds

Variable	Coefficient	T Stat
Social Capital	**0.440*** **	**2.82**
Post Period	-0.294*	-1.73
Insurance	-0.396*	-1.91
Social Capital *Post Period	0.028	0.17
Insurance*Post Period	0.204	0.89
Social Capital *Insurance	0.054	0.28
Social Capital *Post Period*Insurance	**0.590** **	**2.07**
Bond Buyer Index	1.666***	4.78
Log(Issue Size)	1.708***	15.75
Maturity	-0.932***	-16.9
Rating	0.297	1.37
Nonrated Dummy	2.389*	1.83
Competitive Bid	1.037***	5.54
Callable	8.144***	11.19
Bank Qualified	0.835***	3.24
Sinking Fund	7.116***	13.5
Log(Income)	-1.724***	-4.08
Religion Adherence per Capita	-1.312**	-2.06
Religion Diversity Index	-0.179	-1.65
Intercept	87.361	24.55
Year, State Dummies	Included	
N	632,438	
R-Squared	16.84%	

This table shows the regression results of the impact of social capital on volume-weighted trade price in the secondary market. The sample is limited to the trades of general obligation bonds. The model examines whether the impact of social capital on trade price of insured bonds changes after the credit worthiness of insurers deteriorates. To examine this, the study creates a three-way interaction variable of social capital, insurance, and post-downgrade period. Both insurance variable and post-downgrade period are dummy variables. Post period is coded as 1 for the period after downgrade of insurers, otherwise 0, and the variable of insurance is coded as 1 if the bond is insured, otherwise zero. Two-way interactions are included in the model as controls. The regression is clustered by issuance. See Appendix C for variable descriptions in details. N is the sample size of the regression. R-squared represents a goodness of fit measure. Maturity is restricted to positive values. ***,**,* represent significance beyond the 1st, 5th, and 10th percentile levels, respectively.

Results for a Smaller Sample

This study further examines if the main results hold for a smaller sample. The study identifies nine states, which have significant in-state variation in social capital. In-state variation is measured by the standard deviation of social capital of all counties within a particular state. The nine states with the greatest in-state variation in social capital are Kansas, Colorado, North Dakota, South Carolina, Wyoming, Nebraska, Minnesota, Massachusetts and South Dakota. The study examines whether the main results hold under this smaller sample of nine states. The results (untabulated) show that the social capital remains statistically significant in explaining borrowing costs. Holding other variables at the mean, a movement from the 25th percentile of the social capital to the 75th percentile is associated with about a 3.7 percent decrease in bond yields, ceteris paribus[24].

Test on the Robustness of Results based on Actual Reported Data on Social Capital

Social capital data from Rupasingha et al. (2006) are available only on years 1997, 2005, and 2009. In the main analysis, the study linearly interpolated the missing years to extend the sample size. To ensure the results are not driven by the way to interpolate and extrapolate the index, the study examines the main test using only the social capital index on year 2005 and 2009. The results shown in Table 3.13 are consistent with the main analysis. The control variables are untabulated for brevity.

---

[24] 1-(exp(-0.0218*1.266681)/exp(-0.0218*-0.4428091))

Table 3.13: Results are Robust for One Year

	Year=2005		Year=2009	
Variable	Coefficient	T Stat	Coefficient	T Stat
Social Capital	**-0.028*** **	**-3.07**	**-0.052*** **	**-2.4**
Bond-Level Controls	Yes		Yes	
County-Level Controls	Yes		Yes	
Year, State Dummies	Included		Included	
N	66,998		53,454	
R-squared	64.63%		68.65%	

The table presents the coefficient of social capital from the regression analysis when only Rupasingha et al. (2006) data are used. The sample is limited to bonds issued on 2005 and 2009, respectively. Bond-level controls and county-level controls are included in the regression analysis. The controls variables are described in the Appendix C. The standard errors are clustered at the county level.

3.7. Conclusion

This study examines how social capital affects the municipal bond market. High social capital counties are generally associated with greater levels of trust, community-centric values and the propensity to honor obligations. This study assumes that the policy makers within high social capital areas are also more likely to be trustworthy, honest and provide reliable information to investors. These characteristics of high social capital counties are expected to reduce agency concerns and risk, which will reduce bond costs.

The findings show that the cost of borrowing for bonds is lower when municipalities are located in high social capital areas. The high reliability of information in high social capital areas also reduces demand for insurance. The results show that the negative association between bond yield and social capital is especially strong during the post-downgrade period because high social capital acted as a substitute for high quality bond insurance. The argument is also supported by the results on bond prices in the secondary market.

Conclusion

The first part of this dissertation applies methods in audit analytics to improve the efficiency and effectiveness of auditing. Specifically, the first part of the dissertation contains two essays: (1) the application of analytical methods to prioritize exceptions; and (2) the application of the consumer search volume to analytical procedures. The second part of the dissertation analyzes the impact of social capital on the municipal bond market.

The first essay proposes a framework that systematically prioritizes exceptions based on the likelihood of an exception being erroneous. The proposed framework is a semi-automatic system. The framework relies on auditors to initially define and assign confidence levels to the rules and to subsequently investigate the prioritized exceptions. The framework utilizes belief functions to generate suspicion scores for each exception. Then, the framework implements back propagation to increase the system's accuracy in prioritizing erroneous transactions that are higher than normal transactions after each iterative process. Finally, the framework utilizes a rule learner algorithm to generate hidden rules that identify additional characteristics of errors. The proposed framework is evaluated using a simulated experiment. The results from the experiment indicate that the framework is effective in prioritizing exceptions and is able to learn from each iterative run.

The second essay examines whether the consumer search volume can improve the prediction performance and error detection in analytical procedures. The consumer search volume is expected to capture the general level of consumer interest in corporate products or services. This information can be used by auditors to gauge sales and improve analytical procedures. The results indicate that the model incorporating the consumer search volume generates more accurate predictions than conventional models for most consumer-based industries. A simulated experiment was conducted to examine the ability of the model to detect errors. The results indicate that the model incorporating the consumer search volume generally improves error detection for uncoordinated errors and coordinated errors.

The third essay examines the impact of social capital on the cost of debt in the municipal bond market. This study shows that county-level social capital has a strong impact on several aspects of municipal bond offerings. The results indicate that the municipalities in the regions with high social capital issue bonds with lower cost after controlling for bond characteristics. The association between social capital and the cost of debt becomes stronger after the downgrades of bond insurers. The results also indicate that the issuers in the region with high social capital are less likely to purchase the insurance for bonds. These tests indicate that social capital works as a substitute for municipal bond insurance.

BIBLIOGRAPHY

American Institute of Certified Public Accountants (AICPA). 1988. Analytical Procedures. Statement of Auditing Standards No. 56. New York, NY: AICPA.

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

Alles, M. G., Kogan, A., & Vasarhelyi, M. A. (2008). Putting continuous auditing theory into practice: Lessons from two pilot implementations. *Journal of Information Systems, 22*(2), 195-214.

Alles, M., Brennan, G., Kogan, A., & Vasarhelyi, M. A. (2006). Continuous monitoring of business process controls: A pilot implementation of a continuous auditing system at siemens. *International Journal of Accounting Information Systems, 7*(2), 137-161.

Alm, J., & Gomez, J. L. (2008). Social capital and tax morale in Spain. *Economic Analysis and Policy, 38*(1), 73-87.

Ames, D., Brazel, J. F., Jones, K. L., Rich, J. S., & Zimbelman, M. F. (2012). Using nonfinancial measures to improve fraud risk assessments. *Current Issues in Auditing, 6*(1), C28-C34.

Andrei, D., & Hasler, M. (2014). Investor attention and stock market volatility. *Review of Financial Studies,* , hhu059.

Anheier, H. K., Gerhards, J., & Romo, F. P. (1995). Forms of capital and social structure in cultural fields: Examining bourdieu's social topography. *American Journal of Sociology,* , 859-903.

Askitas, N., & Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting. *German Council for Social and Economic Data (RatSWD) Research Notes,* (41)

Baber, W. R., & Gore, A. K. (2008). Consequences of GAAP disclosure regulation: Evidence from municipal debt issues. *The Accounting Review, 83*(3), 565-592.

Baber, W. R., Gore, A. K., Rich, K. T., & Zhang, J. X. (2013). Accounting restatements, governance and municipal debt financing. *Journal of Accounting and Economics, 56*(2), 212-227.

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of german stocks. *Financial Markets and Portfolio Management, 25*(3), 239-264.

Benson, E. D. (1979). The search for information by underwriters and its impact on municipal interest cost. *The Journal of Finance, 34*(4), 871-885.

Benson, E. D., Marks, B. R., & Raman, K. K. (1991). The effect of voluntary GAAP compliance and financial disclosure on governmental borrowing costs. *Journal of Accounting, Auditing & Finance, 6*(3), 303-319.

Bergstresser, D., Cohen, R., & Shenai, S. (2013). Demographic fractionalization and the municipal bond market. *Municipal Finance Journal, 34*(3), 1-38.

Bessembinder, H., Kahle, K. M., Maxwell, W. F., & Xu, D. (2009). Measuring abnormal bond performance. *Review of Financial Studies, 22*(10), 4219-4258.

Blackwell, D. W., & Kidwell, D. S. (1988). An investigation of cost differences between public sales and private placements of debt. *Journal of Financial Economics, 22*(2), 253-278.

Brazel, J. F., Jones, K. L., & Prawitt, D. F. (2014). Auditors' reactions to inconsistencies between financial and nonfinancial measures: The interactive effects of fraud risk assessment and a decision prompt. *Behavioral Research in Accounting, 26*(1), 131-156.

Brazel, J. F., Jones, K. L., & Zimbelman, M. F. (2009). Using nonfinancial measures to assess fraud risk. *Journal of Accounting Research, 47*(5), 1135-1166.

Buonanno, P., Montolio, D., & Vanin, P. (2009). Does social capital reduce crime? *Journal of Law and Economics, 52*(1), 145-170.

Butler, A. W., Fauver, L., & Mortal, S. (2009). Corruption, political connections, and municipal finance. *Review of Financial Studies,* hhp010.

Byrnes, P. E., Al-Awadhi, C. A., Gullvist, B., Brown-Liburd, H., Teeter, C. R., Warren Jr, J. D., et al. (2015). Evolution of auditing: From the traditional approach to the future audit. *Audit Analytics,* 71.

Byrnes, P. E., Ames, C. B., & Vasarhelyi, C. M. (2015). The current state of continuous auditing and continuous monitoring. *Audit Analytics,* 53.

Capeci, J. (1994). Local fiscal policies, default risk, and municipal borrowing costs. *Journal of Public Economics, 53*(1), 73-89.

Chan, D. Y., & Vasarhelyi, M. A. (2011). Innovation and practice of continuous auditing. *International Journal of Accounting Information Systems, 12*(2), 152-160.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16,* 321-357.

Chen, Y., & Leitch, R. A. (1998). The error detection of structural analytical procedures: A simulation study. *Auditing: A Journal of Practice & Theory, 17*(2), 36.

Chen, Y., & Leitch, R. A. (1999). An analysis of the relative power characteristics of analytical procedures. *Auditing: A Journal of Practice & Theory, 18*(2), 35-69.

Choi, H., & Varian, H. (2012). Predicting the present with google trends. *Economic Record, 88*(s1), 2-9.

Chychyla, R., & Kogan, A. (2014). Exploration and exploitation to improve transaction audits. Unpublished working paper. University of Miami, and Rutgers, the State University of New Jersey.

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology, 24*(20), 1-243.

Cohen, J. R., Krishnamoorthy, G., & Wright, A. M. (2000). Evidence on the effect of financial and nonfinancial trends on analytical review. *AUDITING: A Journal of Practice & Theory, 19*(1), 27-48.

Cohen, W. W. (1995). Fast effective rule induction. *Machine Learning-International Workshop then Conference-,* pp. 115-123.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology,* , S95-S120.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance, 66*(5), 1461-1499.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all fears investor sentiment and asset prices. *Review of Financial Studies, 28*(1), 1-32.

deHaan, E., Shevlin, T., & Thornock, J. (2015). Market (in) attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics, 60*(1), 36-55.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *The Annals of Statistics, 28,* 325-339.

Deshmukh, A., & Talluru, L. (1998). A rule-based fuzzy reasoning system for assessing the risk of management fraud. *Intelligent Systems in Accounting, Finance & Management, 7*(4), 223-241.

Dimpfl, T., & Jank, S. (2015). Can internet search queries help to predict stock market volatility? *European Financial Management.*

Domingos, P. (1999). MetaCost: A general method for making classifiers cost-sensitive. *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* pp. 155-164.

Downing, C., & Zhang, F. (2004). Trading activity and price volatility in the municipal bond market. *The Journal of Finance, 59*(2), 899-931.

Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2012). Investor information demand:

Evidence from google searches around earnings announcements. *Journal of Accounting Research, 50*(4), 1001-1040.

Du, R. Y., & Kamakura, W. A. (2012). Quantitative trendspotting. *Journal of Marketing Research, 49*(4), 514-536.

Dugan, M. T., Gentry, J. A., & Shriver, K. A. (1985). The x-11 model-a new analytical review technique for the auditor. *Auditing-a Journal of Practice & Theory, 4*(2), 11-22.

Dzeng, S. C. (1994). A comparison of analytical procedure expectation models using both aggregate and disaggregate data. *AUDITING: A Journal of Practice & Theory, 13*(2), 1.

Fukuyama, F. (1995). Social capital and the global economy. *Foreign Affairs, 74*(5), 89-103.

Furnkranz, J., & Widmer, G. (1994). Incremental reduced error pruning. *International Conference on Machine Learning,* pp. 70-77.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature, 457*(7232), 1012-1014.

Goddard, J., Kita, A., & Wang, Q. (2015). Investor attention and FX market volatility. *Journal of International Financial Markets, Institutions and Money, 38*, 79-96.

Gore, A. K., Sachs, K., & Trzcinka, C. (2004). Financial disclosure and bond insurance. *Journal of Law and Economics, 47*(1), 275-306.

Greif, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy,* 912-950.

Groomer, S. M., & Murthy, U. S. (1989). Continuous auditing of database applications: An embedded audit module approach. *Journal of Information Systems, 3*(2), 53-69.

Guiso, L., Sapienza, P., & Zingales, L. (2000). *The Role of Social Capital in Financial Development,* No. w7563. National bureau of economic research.

Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE, 85*(1), 6-23.

Harris, L. E., & Piwowar, M. S. (2006). Secondary trading costs in the municipal bond market. *The Journal of Finance, 61*(3), 1361-1397.

Hastie, K. L. (1972). Determinants of municipal bond yields. *Journal of Financial and Quantitative Analysis, 7*(03), 1729-1748.

Hirst, D. E., & Koonce, L. (1996). Audit analytical procedures: A field investigation. *Contemporary Accounting Research, 13*(2), 457-486.

Hoitash, R., Kogan, A., & Vasarhelyi, M. A. (2006). Peer-based approach for analytical procedures. *Auditing: A Journal of Practice & Theory, 25*(2), 53-84.

Ingram, R. W., & DeJong, D. V. (1987). The effect of regulation on local government disclosure practices. *Journal of Accounting and Public Policy, 6*(4), 245-270.

Ittner, C. D., & Larcker, D. F. (1998). Are nonfinancial measures leading indicators of financial performance? an analysis of customer satisfaction. *Journal of Accounting Research, 36*, 1-35.

Jain, A. K., & Dubes, R. C. (1988). Algorithms for clustering data. *Englewood Cliffs: Prentice Hall, 1988, 1*

Jha, A. (2013). Financial reports and social capital. *American Accounting Association Annual Meeting, Denver, Colorado,*

Jha, A., & Chen, Y. (2014). Audit fees and social capital. *The Accounting Review, 90*(2), 611-639.

Jiang, F., Li, C. W., & Qian, Y. (2013). Earthly reward to the religious: Religiosity and the cost of public and private debt. Working Paper.

Johnson, T. M., & Grim, B. J. (2013). Religious diversity. *The world's religions in figures* (pp. 93-110) John Wiley & Sons.

Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting, 27*(4), 1116-1127.

Kao, C., & Wu, C. (1994). A re-examination of the impact of credit ratings and economic factors on state bond yields. *Review of Quantitative Finance and Accounting, 4*(1), 59-78.

Kessel, R. (1971). A study of the effects of competition in the tax-exempt bond market. *Journal of Political Economy, 79*(4), 706-738.

Kidwell, D. S., Sorensen, E. H., & Wachowicz, J. M. (1987). Estimating the signaling benefits of debt insurance: The case of municipal bonds. *Journal of Financial and Quantitative Analysis, 22*(03), 299-313.

Kim, Y., & Vasarhelyi, M. A. (2012). A model to detect potentially fraudulent/abnormal wires of an insurance company: An unsupervised rule-based approach. *Journal of Emerging Technologies in Accounting, 9*(1), 95-110.

Kinney, W. R. (1978). ARIMA and regression in analytical review: An empirical test. *The Accounting Review, 53*(1), 48-60.

Kogan, A., Alles, M. G., Vasarhelyi, M. A., & Wu, J. (2014). Design and evaluation of a continuous data level auditing system. *Auditing: A Journal of Practice & Theory, 33*(4), 221-245.

Kogan, A., Sudit, E. F., & Vasarhelyi, M. A. (1999). Continuous online auditing: A program of research. *Journal of Information Systems, 13*(2), 87-103.

Korver, M., & Lucas, P. J. F. (1993). Converting a rule-based expert system into a belief network. *Inform Health Soc Care, 18*(3), 219-241.

Kuenkaikaew, S., & Vasarhelyi, M. A. (2013). The predictive audit framework. *The International Journal of Digital Accounting Research, 13*(19), 37-71.

Kulkarni, G., Kannan, P., & Moe, W. (2012). Using online search data to forecast new product sales. *Decision Support Systems, 52*(3), 604-611.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1997). Legal determinants of external finance. *Journal of Finance,*1131-1150.

Lee, G. H. (2008). Rule-based and case-based reasoning approach for internal audit of bank. *Know.-Based Syst., 21*(2), 140-147.

Lefevre, E., Colot, O., & Vannoorenberghe, P. (2002). Belief function combination and conflict management. *Information Fusion, 3*(2), 149-162.

Leitch, R. A., & Chen, Y. (2003). The effectiveness of expectation models in recognizing error patterns and generating and eliminating hypotheses while conducting analytical procedures. *Auditing: A Journal of Practice & Theory, 22*(2), 147-170.

Lelandais, B., Gardin, I., Mouchard, L., Vera, P., & Ruan, S. (2012). Segmentation of biological target volumes on multi-tracer PET images based on information fusion for achieving dose painting in radiotherapy. *Medical image computing and computer-assisted Intervention–MICCAI 2012* (pp. 545-552) Springer.

Lev, B. (1980). On the use of index models in analytical reviews by auditors. *Journal of Accounting Research, 18*(2), 524-550.

Li, X., Ma, J., Wang, S., & Zhang, X. (2015). How does google search affect trader positions and crude oil prices? *Economic Modelling, 49*, 162-171.

Loebbecke, J. K., & Steinbart, P. J. (1987). An investigation of the use of preliminary analytical review to provide substantive audit evidence. *Auditing-a Journal of Practice & Theory, 6*(2), 74-89.

Maak, T. (2007). Responsible leadership, stakeholder engagement, and the emergence of social capital. *Journal of Business Ethics, 74*(4), 329-343.

Mahoney, J. J., & Mooney, R. J. (1993). Combining connectionist and symbolic learning to refine certainty-factor rule bases. *Connection Science, 5*, 339-364.

Melé, D. (2003). Organizational humanizing cultures: Do they generate social capital? *Journal of Business Ethics, 45*(1-2), 3-14.

Milgram, S., Bickman, L., & Berkowitz, L. (1969). Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology, 13*(2), 79.

Nanda, V., & Singh, R. (2004). Bond insurance: What is special about munis? *The Journal of Finance, 59*(5), 2253-2280.

North, D. C. (1990). *Institutions, institutional change and economic performance* Cambridge university press.

Pastoriza, D., Ariño, M. A., & Ricart, J. E. (2008). Ethical managerial behaviour as an antecedent of organizational social capital. *Journal of Business Ethics, 78*(3), 329-341.

Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Auditing: A Journal of Practice & Theory, 30*(2), 19-50.

Perols, J. L., & Murthy, U. S. (2012). Information fusion in continuous assurance. *Journal of Information Systems, 26*(2), 35-52.

Pietraszek, T. (2004). Using adaptive alert classification to reduce false positives in intrusion detection. *Recent Advances in Intrusion Detection,* pp. 102-124.

Public Company Accounting Oversight Board (PCAOB). PCAOB Standing Advisory Group Meeting: Financial Fraud (September 8–9, 2004).

Public Company Accounting Oversight Board (PCAOB). (2007). Observations on Auditors' Implementation of PCAOB Standards Relating to Auditors' Responsibilities

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy, 6*(1), 65-78.

Putnam, R. D., Leonardi, R., & Nanetti, R. Y. (1994). *Making democracy work: Civic traditions in modern Italy* Princeton university press.

Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing: Explorations in the microstructure of cognition. volume 1-2. foundations.

Rupasingha, A., Goetz, S. J., & Freshwater, D. (2006). The production of social capital in US counties. *The Journal of Socio-Economics, 35*(1), 83-101.

Sasaki, M., & Kita, K. (1998). Rule-based text categorization using hierarchical categories. *Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on, 3.* pp. 2827-2830.

Schultz, P. (2012). The market for new issues of municipal bonds: The roles of transparency and limited access to retail investors. *Journal of Financial Economics, 106*(3), 492-512.

Shafer, G., & Srivastava, R. (1990). The Bayesian and belief-function formalisms a general perspective for auditing. *Readings in uncertain reasoning* (pp. 482-521) Morgan Kaufmann Publishers Inc.

Shafer, G. (1976). *A mathematical theory of evidence* Princeton University Press.

Siganos, A. (2013). Google attention and target price run ups. *International Review of Financial Analysis, 29*, 219-226.

Simonsen, B., & Hill, L. (1998). Municipal bond issuance: Is there evidence of a principal-agent problem? *Public Budgeting & Finance, 18*(4), 71-100.

Simonsen, W., & Robbins, M. D. (1996). Does it make any difference anymore? competitive versus negotiated municipal bond issuance. *Public Administration Review,* 57-64.

Spence, L. J., Schmidpeter, R., & Habisch, A. (2003). Assessing social capital: Small and medium sized enterprises in Germany and the U.K. *Journal of Business Ethics, 47*(1), 17-29.

Srivastava, R. P. (2005). Alternative form of dempster's rule for binary variables. *International Journal of Intelligent Systems, 20*(8), 789-797.

Srivastava, R. P., Gao, L., & Gillett, P. R. (2009). Representation of interrelationships among binary variables under dempster?shafer theory of belief functions. *International Journal of Intelligent Systems, 24*(4), 459-475.

Srivastava, R. P., & Shafer, G. R. (1992). Belief-function formulas for audit risk. *Accounting Review, 67*(2), 249-283.

Styles, A. K., & Tennyson, M. (2007). The accessibility of financial reporting of US municipalities on the internet. *Journal of Public Budgeting, Accounting & Financial Management*, *19*(1), 56.

Thiprungsri, S., & Vasarhelyi, M. A. (2011). Cluster analysis for anomaly detection in accounting data: An audit approach. *The International Journal of Digital Accounting Research, 11*(17), 6.

Towell, G. G., & Shavlik, J. W. (1994). Knowledge-based artificial neural networks. *Artificial Intelligence, 70*(1–2), 119-165.

Trompeter, G., & Wright, A. (2010). The world has Changed—Have analytical procedure practices? *Contemporary Accounting Research, 27*(2), 350-350.

Vasarhelyi, M. A., & Halper, F. B. (1991). The continuous audit of online systems. *Auditing, 10*(1), 110-125.

Vlastakis, N., & Markellos, R. N. (2012). Information demand and stock market volatility. *Journal of Banking & Finance*, *36*(6), 1808-1821.

Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: Survey- based indicators vs. google trends. *Journal of Forecasting*, *30*(6), 565-578.

Walgampaya, C., Kantardzic, M., & Yampolskiy, R. (2010). Real time click fraud prevention using multi-level data fusion. *Proceedings of the World Congress on Engineering and Computer Science,* WCECS 2010, October 20-22, 2010, San Francisco, USA. *, 2186.* (1) pp. 514-519.

Wheeler, S., & Pany, K. (1990). Assessing the performance of analytical procedures: A best case scenario. *The Accounting Review*, *65*(3), 557-577.

White Jr, F. E. (1987). Data fusion lexicon, joint directors of laboratories. *Technical Panel for C, 3.*

Wild, J. J. (1987). The prediction performance of a structural model of accounting numbers. *Journal of Accounting Research*, *25*(1), 139-160.

Wilson, A. C. (1991). Use of regression models as analytical procedures: An empirical investigation of effect of data dispersion on auditor decisions. *Journal of Accounting, Auditing & Finance*, *6*(3), 365-381.

Woolcock, M. (2001). The place of social capital in understanding social and economic outcomes. *Canadian Journal of Policy Research, 2*(1), 11-17.

Wu, L., & Brynjolfsson, E. (2013). The future of prediction: How google searches foreshadow housing prices and sales. *Available at SSRN 2022293,*

Zadrozny, B., Langford, J., & Abe, N. (2003). Cost-sensitive learning by cost-proportionate example weighting. *Data Mining, 2003. ICDM 2003. Third IEEE International Conference on*, pp. 435-442.

APPENDICES

Appendix A

Belief Function Example

The use of belief functions in exception estimation is discussed in this appendix. 'f' is interpreted to imply that the transaction is erroneous, and '$\sim f$' is to imply that the transaction is not erroneous. The entire frame is $\Theta = \{f, \sim f\}$. If transaction *t* violates the rule R_1, the internal auditor may justify a 0.75 degree of belief that the transaction is erroneous, 'f'. However, at the same time, internal auditors have no evidence to indicate that it is not erroneous. This can be presented as follows:

$$\mathbf{m}_{\mathrm{tR}_1}(f) = 0.75,\ \mathbf{m}_{\mathrm{tR}_1}(\sim f) = 0, \text{ and } \mathbf{m}_{\mathrm{tR}_1}(\{f, \sim f\}) = 0.25$$

The internal auditor's judgment about the level of support obtained from rule R_1 can be expressed as:

$$\mathbf{m}_{\mathrm{tR}_1}(f) = \mathrm{r}_1, \quad \text{(A.1)}$$

$$\mathbf{m}_{\mathrm{tR}_1}(\sim f) = 0, \quad \text{(A.2)}$$

$$\mathbf{m}_{\mathrm{tR}_1}(\{f, \sim f\}) = 1 - \mathrm{r}_1, \quad \text{(A.3)}$$

where r_1 is the degree of belief of the auditors that a transaction that violates rule R_1 is treated as erroneous, and $1 - \mathrm{r}_1$ is the ignorance assigned to the entire frame. In this example, r_1 is assumed to be equal to 0.75.

The Belief function from the example is

$$\mathbf{Bel}_{\mathrm{tR}_1}(f) = \mathbf{m}_{\mathrm{tR}_1}(f) = 0.75,$$

$$\mathbf{Bel}_{\mathrm{tR}_1}(\sim f) = \mathbf{m}_{\mathrm{tR}_1}(\sim f) = 0,$$

$$\mathbf{Bel}_{\mathrm{tR}_1}(\{f, \sim f\}) = \mathbf{m}_{\mathrm{tR}_1}(f) + \mathbf{m}_{\mathrm{tR}_1}(\sim f) + \mathbf{m}_{\mathrm{tR}_1}(\{f, \sim f\}) = 1,$$

where $\mathbf{Bel}_{\mathrm{tR}_1}(f) = 0.75$ means that an auditor obtains direct evidence from the rule R_1 that the transactional record is erroneous with 0.75 degree of belief, and $\mathbf{Bel}_{\mathrm{tR}_1}(\sim f) = 0$ means that the auditor does not have direct evidence that the record is not erroneous.

The plausibility function in this example is

$$\mathbf{PL}_{\mathrm{tR}_1}(f) = \mathbf{1} - \mathbf{Bel}_{\mathrm{tR}_1}(\sim f) = 1,$$

$$\mathbf{PL}_{\mathrm{tR}_1}(\sim f) = \mathbf{1} - \mathbf{Bel}_{\mathrm{tR}_1}(f) = 1 - 0.75 = 0.25,$$

where $\mathbf{PL}_{\mathrm{tR}_1}(f)$ implies that there is no degree of belief assigned to '$\sim f$ ', all the probability is assigned to 'f '. $\mathbf{PL}_{\mathrm{tR}_1}(\sim f)$ implies a 0.75 degree of belief is allocated to 'f ', the 0.25 of probability is assigned to '$\sim f$ '.

Srivastava (2005) provides a closed form formula for efficient computation. Following Srivastava (2005), if transaction *t* violates several independent rules, the combined m-value can be represented as:

$$m(f) = 1 - \prod_{i=1}^{n} \big(1 - m_i(f)\big) / K, \tag{A.4}$$

$$m(\sim f) = 1 - \prod_{i=1}^{n} (1 - m_i(\sim f))/K, \tag{A.5}$$

$$m(\{f, \sim f\}) = \prod_{i=1}^{n} m_i(\{f, \sim f\})/K, \tag{A.6}$$

$$\mathrm{K} = \prod_{i=1}^{n}\big(1 - m_i(f)\big) + \prod_{i=1}^{n}(1 - m_i(\sim f)) - \prod_{i=1}^{n} m_i(\{f, \sim f\}), \qquad \text{(A.7)}$$

The suspicion score is given by:

$$Bel_t(f) = m(f) = 1 - \prod_{i=1}^{n}\big(1 - m_i(f)\big)/K \qquad \text{(A.8)}$$

$$\text{Where } \mathrm{K} = \prod_{i=1}^{n}\big(1 - m_i(f)\big) + \prod_{i=1}^{n}(1 - m_i(\sim f)) - \prod_{i=1}^{n} m_i(\{f, \sim f\}) \qquad \text{(A.9)}$$

Back Propagation

All the rules are considered as nodes at the same level. If the internal auditors identify transaction t as erroneous, the real suspicion score for transaction t should be one; otherwise it is zero. The rules that transaction t violates are in the set A_t, and $R_i \in A_t$. Define $\Delta\ r_i$ as the adjustment of the confidence level for rule R_i on the basis of back propagation, and $\Delta_t r_i$ as the adjustment in terms of transaction *t*. Basically, the adjustment for rule R_i is the sum of the adjustments for all the transactions, $\Delta\ r_i = \sum_{t=1}^{n} \Delta_t r_i$. The adjustment of the confidence level r_i for rule R_i from the investigative finding of transaction t is defined as:

$$\Delta_t r_i = \eta \delta_t r_i \qquad \text{(A.10)}$$

where η is the pre-defined adjustment rate, δ_t is the output error, and r_i is the current confidence level for rule R_i. Let E be the overall measure of the error for all the transactions in the investigative sample. The overall measure of the error is the sum-squared error function, $E = \frac{1}{2}\sum_t (T_t - \boldsymbol{Bel}_t(f))^2$. $\mathbf{Bel}_{\mathrm{t}}(f)$ is the assigned suspicion score of transaction t using belief functions and T_t is the real suspicion score of transaction t based on the investigative findings. The error of evaluation δ_t is defined as

$$\delta_t = -\frac{\partial \mathrm{E}}{\partial \mathbf{Bel}_{tR_i}(f)} \tag{A.11}$$

where $\mathbf{Bel}_{tR_i}(f)$ is the confidence level of rule R_i for transaction t. The chain rule is applied to calculate the derivative. It is the product of two parts: the derivative of the transaction's suspicion score with respect to the overall measure of error times the derivative of the rule's belief function with respect to the transaction's suspicion score:

$$\delta_t = -\frac{\partial \mathrm{E}}{\partial \mathbf{Bel}_{tR_i}(f)} = -\frac{\partial \mathrm{E}}{\partial \mathbf{Bel}_{\mathrm{t}}(f)} \frac{\partial \mathbf{Bel}_{\mathrm{t}}(f)}{\partial \mathbf{Bel}_{tR_i}(f)} \tag{A.12}$$

Based on the overall measure of the error $E = \frac{1}{2}\sum_t (T_t - \boldsymbol{Bel}_t(f))^2$, it will be

$$\frac{\partial \mathrm{E}}{\partial \mathbf{Bel}_{\mathrm{t}}(f)} = -(T_t - \boldsymbol{Bel}_t(f)) \tag{A.13}$$

The rules in the study can either be affirmative evidence or negative evidence. Although the formula for the suspicion score for transactions will be the same, the calculation for back propagation is different. The calculation is derived respectively below.

Affirmative Evidence

We know $\mathbf{Bel}_{tR_i}(f) = \mathrm{m}_i(f)$, where r_i is the confidence level of rule R_i.

$$\frac{\partial \mathbf{Bel}_{\mathrm{t}}(f)}{\partial \mathbf{Bel}_{tR_i}(f)} = \frac{\prod_{j=1}^{n}\left(1 - m_j(f)\right) m_i(f)}{K^2 * (1 - m_i(f))} [\mathrm{K} + \prod_{j=1}^{n}\left(1 - m_j(f)\right)]$$

$$\delta_t = -\frac{\partial \mathrm{E}}{\partial \mathbf{Bel}_{\mathrm{t}}(f)} \frac{\partial \mathbf{Bel}_{\mathrm{t}}(f)}{\partial \mathbf{Bel}_{tR_i}(f)}$$

$$= (T_t - \boldsymbol{Bel}_t(f)) \frac{\prod_{j=1}^{n}\left(1-m_j(f)\right)m_i(f)}{K^2 * (1-m_i(f))} [\mathrm{K} + \prod_{j=1}^{n}\left(1 - m_j(f)\right)] \quad \text{(A.14)}$$

The adjustment of the confidence level of rule R_i based on the investigative findings on transaction *t* is:

If transaction t is erroneous, the revision for the confidence level of R_i is

$$\Delta_t r_i = \left(1 - \mathbf{Bel}_{\mathrm{t}}(f)\right) * \frac{\prod_{j=1}^{n}\left(1-m_j(f)\right)m_i(f)}{K^2 * (1-m_i(f))} [\mathrm{K} + \prod_{j=1}^{n}\left(1 - m_j(f)\right)] * r_i \quad \text{(A.15)}$$

If transaction t is not erroneous, the revision for the confidence level of the R_i is

$$\Delta_t r_i = \left(0 - \mathbf{Bel}_{\mathrm{t}}(f)\right) * \frac{\prod_{j=1}^{n}\left(1-m_j(f)\right)m_i(f)}{K^2 * (1-m_i(f))} [\mathrm{K} + \prod_{j=1}^{n}\left(1 - m_j(f)\right)] * r_i \quad \text{(A.16)}$$

where K is the renormalization that was defined in stage 2. n includes these rules that the transaction violates.

Negative Evidence

$$\frac{\partial \mathbf{Bel}_{\mathrm{t}}(f)}{\partial \mathbf{Bel}_{tR_i}(f)} = -\frac{\prod_{j=1}^{n}\left(1 - m_j(\sim f)\right) * \prod_{j=1}^{n}\left(1 - m_j(f)\right) m_i(\sim f)}{K^2(1 - m_i(\sim f))}$$

$$\delta_t = -\frac{\partial \mathrm{E}}{\partial \mathbf{Bel}_{\mathrm{t}}(f)} \frac{\partial \mathbf{Bel}_{\mathrm{t}}(f)}{\partial \mathbf{Bel}_{tR_i}(f)} = (T_t - \boldsymbol{Bel}_t(f))[-\frac{\prod_{j=1}^{n}\left(1-m_j(\sim f)\right)*\prod_{j=1}^{n}\left(1-m_j(f)\right)m_i(\sim f)}{K^2\left(1-m_i(\sim f)\right)}]$$

(A.17)

The adjustment of the confidence level of rule R_i based on the investigative findings on transaction *t* is:

If transaction t is erroneous, the revision for the confidence level of R_i is

$$\Delta_t r_i = \left(1 - \mathbf{Bel}_{\mathrm{t}}(f)\right)[-\frac{\prod_{j=1}^{n}\left(1-m_j(\sim f)\right)*\prod_{j=1}^{n}\left(1-m_j(f)\right)m_i(\sim f)}{K^2\left(1-m_i(\sim f)\right)}] * r_i \qquad \text{(A.18)}$$

If transaction t is not erroneous, the revision for the confidence level of R_i is

$$\Delta_t r_i = \left(0 - \mathbf{Bel}_{\mathrm{t}}(f)\right)[-\frac{\prod_{j=1}^{n}\left(1-m_j(\sim f)\right)*\prod_{j=1}^{n}\left(1-m_j(f)\right)m_i(\sim f)}{K^2\left(1-m_i(\sim f)\right)}] * r_i \qquad \text{(A.19)}$$

The new confidence level of rule R_i incorporates the prior confidence level and the aggregative adjustment for all transactions.

$$r_i{}' = r_i + \eta \; \Delta \; r_i \qquad \text{(A.20)}$$

where η is the pre-defined adjustment rate, and $\Delta \; r_i = \sum_{t=1}^{n} \Delta_t r_i$.

Appendix B

Artificial Errors Instances Generation

K-means clustering is used to artificially label erroneous accounts payable transactions. The process consists of three steps; 1) transformation of accounts payable data, 2) clustering of transactions, and 3) labeling of anomalies (errors). These three steps are shown in Figure B.1 and discussed further below.

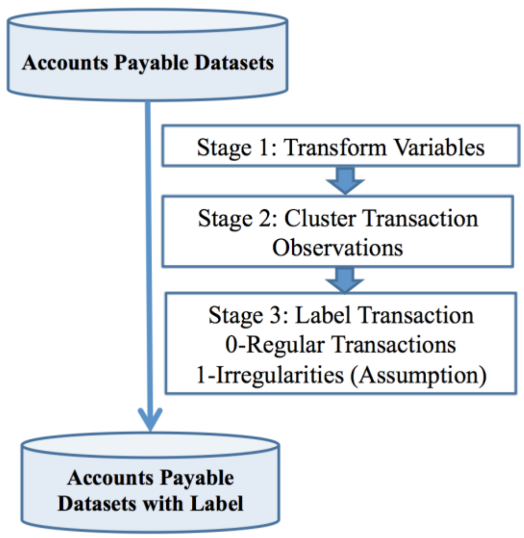

Figure B. 1: Generation of Artificial Errors Instances

Step 1: Transformation of Accounts Payable Data

There are two types of variables used in clustering; 1) the original variables from the accounts payable data and 2) the transformed variables. Table 8 presents these variables. Transformed variables are used to provide the clustering system with additional variables to improve and support its performance (Jain and Dubes 1988). These transformed variables are derived from and relate to the expert rules. For example, one expert rule is to find those transactions that occur on weekends. The transformed variable is called Payment_On_Weekend as shown in Table B.1. It is a dummy variable, taking the value of 1 when the transaction occurs on a weekend and 0 otherwise. In this study, there are seven original variables used from the accounts payable data and twenty transformed variables.

Table B.1: Clustering Variables

Variables	Variable Type	Description
Invoice_Date	Transaction	Invoice Date
Tax_Amount	Transaction	Tax Amount
Goods_Amount	Transaction	Goods Amount
Invoice_Type	Transaction	Invoice Type
Full_Pay_Status	Transaction	Full Pay Status
General Ledge Account	Transaction	General Ledge Account
Payment ID	Transaction	Payment ID
Miss_Date	Expert Rule Related	Transaction with missing disbursement date
Miss_Invoice_Type	Expert Rule Related	Transactions with missing invoice type
Invalid_GL	Expert Rule Related	Transactions with invalid or missing GL account number
Miss_Invoice_No	Expert Rule Related	Transactions with missing invoice number
RoundAmount_Line	Expert Rule Related	Found the lines with round amounts (e.g. 100; 1,000; 10,000 etc.)
RoundAmount_Invoice	Expert Rule Related	Found the invoices with round amounts (e.g. 100; 1,000; 10,000 etc.)
Fraud_KeywordsSearch	Expert Rule Related	Identify disbursements containing one of the keywords in Fraud category
FCPA_KeywordsSearch	Expert Rule Related	Identify disbursements containing one of the keywords in FCPA category
Outlier_Disbursement	Expert Rule Related	Summarized the disbursement by vendor ID, and identified any outliers
Outlier_GL	Expert Rule Related	Summarized the disbursements file by GL Account Number
Invalid_Vendor	Expert Rule Related	Identify disbursements to vendors not found in the vendor master file
Duplicate	Expert Rule Related	Search for duplicate disbursements
Payment_Due_Date	Expert Rule Related	Analyzed payment dates in reference to due dates
Payment_Invoice_Date	Expert Rule Related	Analyzed payment dates in reference to invoice dates
GAP_Voucher_No	Expert Rule Related	Identify gaps in the voucher sequence
Payment_Negative_Amount	Expert Rule Related	Searched for any negative disbursement amounts
Payment_Zero_Amount	Expert Rule	Searched for zero disbursement amounts

	Related	
Payment_On_Weekend	Expert Rule Related	Payments on the weekend
Payment_On_Holiday	Expert Rule Related	Payments on holidays
MultiInvoice_Per_Day	Expert Rule Related	Multiple invoices on one day to the same vendor

The Expectation Maximization (EM) algorithm in WEKA[25] is used to cluster the transactional records. The EM algorithm in WEKA uses the log likelihood as the criterion to determine the optimal number of clusters. A larger value for the log likelihood indicates a better fit. The log likelihoods for different numbers of clusters are compared. In this study, the optimal number of clusters was fourteen.

The EM algorithm assigns each instance a probability over the clusters. Such a probability indicates the likelihood of an instance belonging to a certain cluster. The results generated by WEKA are presented in Table B.2. The second column shows the number of transactions that were located in a certain cluster. The value in the parentheses shows its percentage in the whole population. In the sensitivity test, the second optimal number of clusters was 17 and was used after comparing with the different number of clusters.

[25] http://www.cs.waikato.ac.nz/ml/weka/

Table B.2: Clustering Results

Cluster Number	Number of Transactions	Percentage
1	666	0.007
2	10,176	0.113
3	5,809	0.065
4	11,215	0.125
5	2,869	0.032
6	12,576	0.140
7	7,217	0.080
8	10,827	0.121
9	7,673	0.086
10	3,537	0.039
11	2,327	0.026
12	3,644	0.041
13	1,641	0.018
14	9,535	0.106

Step 3: Labeling of Anomalies

Two methods are utilized to determine the anomalies in the accounts payable data: the size of clusters and the distance between the observation and the nearest clustered centroid (Thiprungsri and Vasarhelyi 2011). The former method is to treat transactions in large and dense clusters as regular transactions and transactions in small or sparse clusters as anomalies. These types of anomalies are often called cluster-based anomalies. Suppose three clusters are generated: the first cluster consists of 1000 transactions, the second cluster consists of 2000 transactions, and the third cluster consists of 10 transactions. All the transactions in the third cluster are treated as cluster-based anomalies. The latter method is to treat the transactions far away from its nearest clustered centroid as anomalies. The distance between the observation and its nearest cluster in the EM algorithm is represented as the probability of the observation belonging to the cluster. Those transactions deemed to be anomalies are labeled as erroneous transactions. In the previous example, two variables were used for clustering, transaction date and transaction amount. Most of the transactions in the first cluster occurred during year 2013 and valued around $1,000. Among them, there are twenty transactions occurring during year 2014 and valued around $10,000. Those transactions are treated as anomalies.

The results of clustering are analyzed below (Table 9). Over 60% of transactions are grouped into cluster 2, cluster 4, cluster 6, cluster 8, and cluster 14. These clusters are large and dense. On the contrary, there is one cluster with less than 1% of the observations. That cluster is cluster 1 which includes 666 observations. All members in the cluster are considered as cluster-based anomalies. The probability of each member is further analyzed to determine its membership in a cluster. It is reasonable to not consider an observation with probability of less than 0.5 as a member of its cluster. To obtain enough observations as anomalies, those members with probabilities of less than 0.6 are treated as anomalies. This threshold for the probability is set arbitrarily. In the study, the amount of observations that have a low probability of being a member of a cluster is 899. Therefore, the total number of anomalies identified using clustering is 1,565 (1.74%).

There are five alternative methods to deal with an unbalanced dataset issue; 1) oversampling (Chawla, Bowyer, Hall, and Kegelmeyer 2002), 2) under-sampling (Chawla et al. 2002; Perols 2011), 3) a combination of oversampling and under-sampling (Chawla et al. 2002), 4) cost-sensitive classifier (Zadrozny, Langford, and Abe 2003), and 5) meta-cost classifier (Domingos 1999). Table B.3 provides the illustration for the five alternatives. Each alternative method is applied on the training subset to deal with the imbalance issue.

Table B.3: Illustrations of the Five Alternatives

Alternative	Methodology Description	Details
Original	N/A	Work as a comparison.
Under-sampling	Balance the dataset through under-sampling the majority class.	In the experiment, the under-sampling ratio of the majority class to the minority class varies from 1.0 to 1.9. The under-sampling with the ratio of 1.1 performs superior to the under-sampling with other ratios.
Oversampling	Oversample the minority class by creating some synthetic observations in the minority class.	Vary the oversampling ratio of the minority class from 50% to 200%. The data with the oversampling ratio 250% is superior to others.
Combination of oversampling and under-sampling	Oversample the minority class to a specified degree, and then under-sample the majority class equal to the minority class.	Compare the combinations of over-sampling and under-sampling with different ratios. When we oversample the minority class to 50 percent and under-sample the majority class with 110 percent ratio to the minority class in combination, this combination performs better than others.
Cost-sensitive classifier	Introduce distinct weights to observations using a cost matrix that represents the cost of misclassification.	The study tests this method with the relative cost ratio of misclassifying errors versus misclassifying normal transactions. This ratio ranges from 3:1 to 40:1 in the test. The relative cost ratio 35:1 works better than others.
Meta-cost classifier:	Produces multiple replicas of the data and learns the classifier on each of the replicas. The classification of each observation considers both the vote aggregated from these classifiers and the cost of misclassification.	The relative cost of the misclassification of an error to the misclassification of one regular transaction ranges from 2:1 to 65:1. The performance of the type of classifiers has not changed since the relative cost ratio is larger than the ratio value of 13:1.

A confusion matrix is used to evaluate the performance of the five alternatives (Table B.4). There is a high cost of misidentifying erroneous transactions. A lower false positive rate indicates a more effective method in dealing with the imbalance issue in this experiment. False positive rate is calculated as the percentage of errors that is incorrectly identified as normal transactions. Since the performance of each method varies with the value of its parameters, the results of each method with different parameters are compared, and the one that outperforms the others is selected. The results of each method with the appropriate parameters are presented in Table B.5. According to these results, the combination method is superior to the other four alternatives. Therefore, the combination method was selected to deal with the imbalance issue in the study. However, the combination method may not be the preferred approach in all situations since the method depends on the dataset.

Table B.4: Confusion Matrix Definitions

	Predicted as Normal Transactions	Predicted as Errors
Normal Transactions	TP [a]	FN [c]
Errors	FP [b]	TN [d]

[a] True Positive (TP) is the number of regular transactions correctly identified.
[b] False Positive (FP) is the number of erroneous transactions incorrectly identified as regular transactions.
[c] False Negative (FN) is the number of regular transactions incorrectly identified as erroneous transactions.
[d] True Negative (TN) is the number of erroneous transactions correctly identified.

Table B.5: Confusion Matrix - Comparisons of Methods Dealing with an Unbalanced Dataset

	Original	Under-sampling	Oversampling	Combination	Cost-Sensitive Classifier	MetaCost Classifier
FP rate [a]	0.1832	0.079	0.628	0.070	0.087	0.828
FN rate [b]	0.0003	0.143	0.002	0.150	0.075	0.002

[a] FP rate: measures the proportion of erroneous observations which are incorrectly identified as regular transactions.

[b] FN rate: measures the proportion of regular observations that are incorrectly identified as erroneous observations.

Appendix C

Table C.1: Variables Measurement

Variables Name	Description and Measurement
Main Dependent Variables:	
Bond Yield	Yield to maturity at the time of issuance based on the coupon and any discount or premium to par value at the time of sale
Yield Spread	The difference between the yields on a municipal bond and a U.S. treasury bond. A municipal bond is compared against treasuries by duration.
Main Research Variables	
Social Capital	The measure of the social capital is at the county level. The construction of the variable follows Rupasingha, Goetz, and Freshwater (2006).
Bond-Level Control Variables	
Bond Size	The principal amount of the maturity's original offering individual bond. The natural logarithm of this variable is used in the models.
Issue Size	Natural logarithm of the total par value of all bond issues in a deal. There are several bonds with different maturities, offering yield, coupon, etc.
Maturity	The maturity of the bond, measured in years.
Bond Buyer Index	The 20-Bond Index from the bond buyer consists of 20 general obligation bonds that mature in 20 years.
Bank Qualified	An indicator variable that takes a value of one for a bond that banks can deduct the interest expense for the purchase or carry of these obligations.
Insurance	An indicator variable for the bond having insurance. The variable takes a value of one for bonds with insurance and zero otherwise.
GO Bond	An indicator variable that takes a value of one for bonds those are general obligation bonds for a county and zero otherwise.
Competitive Bid	An indicator variable that takes a value of one for bonds for which the underwriter is engaged through a competitive offer and zero otherwise.
Callable	An indicator variable that Issuer is permitted or required to redeem the bond between the transaction date and maturity.
Bond Rating	A numerical categorization of the bond’s credit rating assigned by rating agencies.
No Rating Dummy	This variable is created to takes a value of 1 to capture the fact that the bond is rated, otherwise unrated.
County-Level Control Variables:	

Income	The income per capita in a county is deflated by the Consumer Price Index. The natural logarithm of this variable is used in the models.
Religion Adherents Per Capita	The rate of religious adherents per capita
Religion Diversity Index	The index is to measure the levels of religious diversity. It was developed by Brian J. Grim on the basis of the Herfindahl index used in economics. The higher score of religion diversity index indicates higher diversity of religion.

Table C.2: Classification of Bond Ratings

S&P or Fitch Ratings	Moody's Ratings	Numerical Code
Not rated	Not rated	-1
Below BBB+	Below Baa1	2
A-,BBB+	Baa1, A3	3
A+,A	A, A1, A2	4
AA, AA−	Aa2, AA3	5
AA+, AAA	Aa, Aa1, Aaa	6

This table lists the numerical codes associated with the ratings assigned by Moody's, S&P and Fitch.