THREE ESSAYS ON EMERGING TECHNOLOGIES IN ACCOUNTING

BY

FEIQI HUANG

A dissertation submitted to the

Graduate School - Newark

Rutgers, The State University of New Jersey

In partial fulfillment of requirements

For the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Dr. Miklos A. Vasarhelyi

And approved by

____

Dr. Miklos A. Vasarhelyi

____

Dr. Won Gyun No

____

Dr. Alexander Kogan

____

Dr. Tawei Wang

Newark, New Jersey

January 2019

ABSTRACT OF THE DISSERTATION

THREE ESSAYS ON EMERGING TECHNOLOGIES IN ACCOUNTING

BY FEIQI HUANG

Dissertation Director: Dr. Miklos A. Vasarhelyi

This dissertation consists of three essays that examine the effects of emerging technologies in accounting. The first essay examines whether firms abuse XBRL extension elements to increase the complexity of their mandatory filings in interactive data format. Using the ratio of extension elements to total elements in XBRL 10-K filings as the measure of XBRL complexity, this study finds that firms' XBRL filings are more complex when the firms are performing poorly, suggesting that managers use extension elements strategically to increase XBRL complexity and obfuscate XBRL-tagged financial information.

The second essay investigates whether the adoption of the XBRL affects firms' capital investment efficiency due to the increased information processing efficiency. The findings of this essay suggest that the adoption of XBRL reduces the levels of abnormal investments. Additionally, the benefits of XBRL mandate on investment efficiency are more evident for firms with weaker external monitoring, severer environmental uncertainty, and more readable financial reports.

The third essay introduces robotic process automation (RPA) to the auditing area. A framework is proposed to apply RPA to audit procedures in order to free auditors from doing repetitive and low-judgment audit tasks and enable them to focus on audit tasks that

require professional judgment. This essay also demonstrates the feasibility of RPA by implementing a pilot project that applies RPA to the confirmation process.

In conclusion, this dissertation examines the effects of XBRL on financial reporting strategy and managers' investment decision, proposes a framework to apply RPA to automate labor-intensive, well-defined, and repetitive audit procedures, and demonstrate the feasibility in the audit practice.

ACKNOWLEDGMENTS

First and foremost I would like to express my deepest gratitude to Dr. Miklos A. Vasarhelyi, who has been not only a supportive academic advisor but also a truly dedicated mentor. I am grateful to his continuous support of my doctoral study, to his encouragement during tough times in my Ph.D. pursuit, and to his insightful advice about my academic career. The enthusiasm, persistence, and integrity he has for research were contagious and motivational for me.

I would also like to extend my sincere thanks to Dr. Won Gyun No, who is generous giving in time and patient to encourage me, to share his stories with me, and to contribute valuable advice to my research. I am particularly thankful to Dr. Tawei (David) Wang, who guides me step-by-step in attempting to get my first publication. I very much appreciate Dr. Alex Kogan for serving on my dissertation committee and for his time and insightful comments.

Special thanks to my friends and colleagues at Rutgers and CARLab, especially He Li, Yang Liu, Cheng Yin, Yunsen Wang, Tiffany Chiu, Qiao Li, Zhaokai Yan, Yue Liu, Ting Sun, Jiahua Zhou, Lu Zhang, and Xin Cheng. You guys made my time at Rutgers enjoyable.

Lastly, I acknowledge the highest appreciation to my family for their constant support and unconditional love. For my parents who raised me with love and supported me in all my pursuits. For my puppy Pepper who pushes me to get up early no matter what is the day of the week and brings me endless happiness. This last word of acknowledgment

is saved for my beloved wife Jue Wang, who keeps me motivated to study since I know that my only excuse for not cooking or walking Pepper is “I am doing research!”

Feiqi Huang
Rutgers University
January 2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Advances in technologies have almost changed every aspect of the world and never before in history has that change occurred so fast. The emerging technologies have reengineered business processes, redefined business environment, and remodeled many aspects in business. The field of accounting is undergoing a fundamental change as well. Investors, managers, regulators, and auditors are all facing new opportunities and challenges. This dissertation studies the impact of emerging technologies in accounting and examines their effects on the financial reporting process and the audit practice. Specifically, I focus on two technologies: the eXtensible Business Reporting Language (XBRL) and Robotic Process Automation (RPA).

XBRL is a global open standard for preparing, publishing, exchanging, and consuming financial information. In XBRL, financial facts in financial statements are tagged using pre-defined machine-readable elements. On January 30, 2009, the Securities and Exchange Commission (SEC) issued final rules (i.e., 33-9002) that mandate the use of XBRL for financial reporting (SEC 2009). Since the XBRL adoption, prior literature has extensively documented the benefits of XBRL adoption such as increased transparency, reduced information asymmetry, and improved accessibility (e.g., Blankespoor et al. 2014; Kim et al. 2012; Liu et al. 2014). The first two essays extend this line of literature and study the managers' response to the adoption of XBRL in financial reporting process.

The first essay examines whether firms use XBRL elements to increase the complexity of their mandatory filings in interactive data format. Using the ratio of extension elements to total elements in XBRL 10-K filings as the measure of XBRL complexity, this study finds that firms' XBRL filings are more complex when the firms are performing poorly. The analysis of the relation between firms' future performance and XBRL complexity shows that complex XBRL filings are associated with less (more) persistent positive (negative) earnings. The evidence further reveals that this effect is more pronounced when firms are inherently more complex. Collectively, the results suggest that managers use extension elements strategically to increase XBRL complexity and obfuscate XBRL-tagged financial information.

The second essay examines the effect of the adoption of XBRL, and investigate whether the reduced information processing cost affects capital investment efficiency. The findings of this study reveal that the adoption of XBRL reduces the level of abnormal investments. To investigate potential moderating factors that may magnify or mitigate the benefits of XBRL on the improvement in investment efficiency, several analyses are conducted and the results show that the benefits of XBRL mandate on investment efficiency are more evident for firms with weaker external monitoring, severer environmental uncertainty, and more readable financial reports. The additional analyses show that results are robust to difference-in-difference (DID) setting, change research design and a non-capital investment setting. Finally, an increasing pattern of the effect on investment efficiency is identified, supporting that investors face a learning curve in understanding XBRL (Du et al. 2013).

Traditional audit procedures are labor-intensive and time-consuming (Chan and Vasarhelyi 2011). To free human auditors from doing repetitive and low-judgment audit tasks and help them to focus on procedures requiring professional judgment, prior literature has proposed for decades that labor-intensive audit tasks be replaced with automation (e.g., Vasarhelyi 1984; Vasarhelyi and Halper 1991). Although technology has had a significant impact on improving audit efficiency, integration across multiple systems or applications is performed mainly by auditors, meaning that the actual external audit is still labor-intensive (Srinivasan 2016).

Recently, practitioners have been interested in rethinking their process in the line of automation and taking advantage of advanced automation technologies such as robotic process automation (RPA). RPA is a methodology that performs routine business processes by automating the way people interact, with multiple applications or analyses through a user interface and also by following simple rules to make decisions (Deloitte 2017). Because of its low implementation cost and high potential benefits, RPA has been widely adopted in many industries. Even though the industry has observed the benefits of RPA, its applications in auditing practice are still unexplored. Additionally, many audit tasks are well-defined, highly repetitive, and predictable; for example, extracting exogenous information (confirmations from the electronic platform and customer reviews from social media) and matching information from multiple systems, which are multi-step tasks across multiple systems, are the ideal candidates for RPA (IRPA 2015). With the improved processing power of RPA, the scale of audit procedures can be increased and auditors will be able to focus on tasks that require professional judgment and higher order thinking skills, thereby enhancing audit quality. In the third essay, we propose a framework to apply RPA in auditing and demonstrate its feasibility by implementing a pilot project for the confirmation process.

CHAPTER 2: LITERATURE REVIEW

2.1. eXtensible Business Reporting Language (XBRL)

On January 30, 2009, the Securities and Exchange Commission (SEC) issued final rules (i.e., 33-9002) that mandate the use of an interactive data format known as the eXtensible Business Reporting Language (XBRL) for financial reporting (SEC 2009). XBRL is a global open standard for preparing, publishing, exchanging, and consuming financial information. In XBRL, financial facts in financial statements are tagged using pre-defined machine-readable elements. An electronic dictionary of such elements is called taxonomy and defines each element that represents specific financial concepts (e.g., net income) and the relationships among elements (i.e., how a financial concept is related to other concepts) (AICPA 2007). A taxonomy is generally developed by regulatory bodies such as the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) and known as standard (or official) taxonomy. By tagging each fact in financial statements using taxonomies, XBRL allows machines to understand what a tagged number represents based on the element used (i.e., which financial concept) and how it relates to other numbers. Hence, XBRL not only makes financial data more accessible and reliable but also allows analysts, investors, regulators, and related parties to handle financial data faster, easier, and cheaper, and thus improves their analysis and decision-making (Dong et al. 2016; XBRL.US 2014).

Since the SEC mandated public firms to use XBRL in their financial reporting in 2009, prior literature has extensively documented the benefits of XBRL adoption such as increased transparency, reduced information asymmetry, and improved accessibility (e.g., Blankespoor et al. 2014; Kim et al. 2012; Liu et al. 2014). For example, Cong et al. (2014)

and Kim et al. (2012) find a significant decrease in information risk and information asymmetry after XBRL adoption, implying an improved information quality. Dong et al. (2016) provide evidence that consistent with the SEC's statement, XBRL adoption helps market participants translate more firm-specific information into stock prices. Additionally, due to the improved information accessibility resulting from XBRL adoption, analyst coverage and the timeliness and accuracy of analyst forecasts are significantly enhanced (Liu et al. 2014). Furthermore, increased external monitoring attributed to XBRL adoption may also affect managers' behaviors. Chen et al. (2016) show that the extent of corporate tax avoidance decreases significantly after the adoption of XBRL for financial reporting, and this pattern is more pronounced for firms with lower levels of institutional ownership and analyst coverage.

More relevant to my dissertation, the adoption of XBRL enhances investors' information-processing capacity as well. The mandate of XBRL filings is to give small investors more accessible financial information in a user-friendly and less costly search-facilitating information environment (SEC 2009). Compared to sophisticated investors, small investors generally have fewer resources and limited ability to process information (Blankespoor et al. 2014), which constrains their decision-making ability. XBRL-enhanced search engines enable investors to view financial information with similar tags simultaneously, improving investors' analytical capabilities (XBRL.US, 2009). Since the adoption of XBRL makes it simpler to search, extract, and confirm firm-specific information, it should be much easier for small investors to conduct basic analysis, such as comparing financial ratios among competitors, evaluating suggestions from other information channels, and generating their own opinions based on the firm-specific information. Consistent with this view, Hodge et al. (2004) conduct behavioral experiments and show that non-professional users are likely to benefit from search-facilitating technologies like XBRL in analyzing financial reports. In addition, the adoption of XBRL brings more opportunities for smaller investors to analyze firm performance and to play a monitoring role in corporate governance. They can execute the power of shareholders at a lower cost by utilizing the tool of XBRL. Small investors' ability to monitor the wrong-doings of managers puts "additional perceived pressures" on managerial decisions, due to managers' career concerns, such as the fear of involuntary replacement and the desire to influence the markets' perception of their ability positively (Holmstrom, 1982; Ali and Zhang, 2015). XBRL's enhanced monitoring functionalities may also effectively reduce suboptimal managerial behaviors with regard to investment efficiency because of the increased probability that managerial misconducts will be detected.

The adoption of XBRL not only favors smaller investors, but also facilitates sophisticated users' ability to access, extract, and analyze firm performance data more efficiently and effectively. Through more efficient and effective processing of financial reporting, sophisticated information users, such as institutional investors and analysts, are able to leverage their superior knowledge to obtain greater benefits from XBRL and enhance their information advantages (Blankespoor et al., 2014). In the pre-XBRL era, sophisticated information users bore the cost of information mining through self-supported agents (applications, software, and other programming-based intelligence macros). Searching, extracting, and formalizing data from complex, diversified financial reports takes resources away from analyzing the disclosed financial details. The XBRL mandate eliminates the need to convert financial information into machine-readable records, giving sophisticated users more time to conduct the value-added analysis. For instance, Bloomberg consumes XBRL data to fast-track company financials to analysts and has increased the usage every year including the footnotes which may not be captured in the past (Efthimides 2017). Liu et al. (2014) document that XBRL adoption increases analyst coverage, and improves the timeliness and enhances the accuracy of analysts' forecasts. Consequently, by shifting more resources from tedious information collection to information analysis, larger investors may improve their ability to evaluate firm and management performance. Better analyzing power enables investors to detect managerial opportunism and possible misalignments in an effective and timely manner in the post-XBRL era. Thus, the XBRL mandate may curb suboptimal investment decisions made by managers.

According to SEC (2009), firms are required to tag each amount (i.e., monetary value, percentage, and number) in their financial statements using an element in the U.S. GAAP Financial Reporting Taxonomy (hereafter U.S. GAAP taxonomy), which is created and managed by FASB as the standard (or official) taxonomies (xbrl.fasb.org). For example, "Cash and Cash Equivalents" in the balance sheet could be tagged using the standard "CashAndCashEquivalentsAtCarryingValue" element in the U.S. GAAP taxonomy. Under the SEC mandate, firms can also create extension taxonomies by defining their own elements, called extension elements when appropriate elements do not exist in the official U.S. GAAP taxonomy. Although creating extension elements is likely to enhance reporting flexibility (SEC 2009), the use of extension elements requires manual interpretation of tagged data, reduces machine readability, and hinders cross-firm comparability (Boritz and No 2009). Practitioners report concerns regarding the unnecessary use of extension elements (XBRL.US 2010). The SEC has also issued series of announcements to warn the use of unnecessary extensions (SEC 2011, 2010, 2014). Consistent with this concern, researchers find initial evidence that managers may overuse extensions (Debreceny et al. 2011), and the use of extension elements may impair market efficiency. Dhole et al. (2015) find that the comparability of financial statement is reduced by the use of extension elements. Additionally, analysts suffer from increased complexity, which in turn leads to lower forecast accuracy and greater dispersion (Kirk et al. 2016).

Since the XBRL tagging process (i.e., choosing appropriate elements and creating extension elements for financial facts) often involves significant management discretion (Kirk et al. 2016), what leads to the overuse of extension elements has been a focus of existing literature. On the one hand, Guragai et al. (2014) point out that together with the individuals' inexperience interpreting XBRL filings, managers have an opportunity to mispresent financial information with a lower chance of detection by increasing the complexity of XBRL filings. Furthermore, Hoitash and Hoitash (2017) suspect that managers' attempt at obfuscating information could lead to the overuse of extension elements since they observe that extension elements are associated with discretionary accruals. On the other hand, Scherr and Ditter (2017) do not find an association between the use of extension elements and earnings quality and suggest that filing complexity and the cost of compliance are the main determinants of the deviation from using standard elements. Accordingly, more research is needed to provide insights into the debate of whether managers strategically overuse extension elements to introduce the complexity of XBRL filings (henceforth XBRL complexity) and obfuscate XBRL-tagged financial information.

2.2. Robotic Process Automation (RPA)

The idea behind RPA is not new; it is traditional automation in terms of assembly line technology (Moffitt et al. 2018). RPA is defined as "*a preconfigured software instance that uses business rules and predefined activity choreography to complete the autonomous execution of a combination of processes, activities, transactions, and tasks in one or more unrelated software systems to deliver a result or service with human exception management*" (IEEE Corporate Advisory Group 2017, p11). The RPA works as an overlay for existing IT systems. An example of the RPA process is the retrieval of information from one system and entering the same information into another system or activating another system function. Unlike some traditional IT implementation and business reengineering that changes the existing systems, RPA tries not to disturb underlying IT systems and only replaces the existing manual process with the automated process through a presentation layer (IRPA 2016). Therefore, compared with major IT platform updates, the burdens of RPA implementation (costs, timelines, and risks) are relatively insignificant (EY 2016b).

RPA tools help businesses improve the efficiency of processes and the effectiveness of services. First, replacing the human workforce reduces the cost and processing time for high-frequency tasks. The running cost of an RPA software is around one-ninth that of employing a human being, and RPA "robots" can work 24 hours a day and 7 days a week (Burgess 2016). Secondly, the accuracy of the business process is improved. As long as an RPA tool is properly programmed, there is no need to worry that software robots will make the mistakes that human beings might (IRPA 2015). Finally, RPA offers flexibility and scalability. Once a process has been executed by a software robot, it can be scheduled for a particular time. In addition, the RPA robot is capable of performing many types of processes and can be quickly reassigned to other processes (Deloitte 2017).

Because of its low implementation cost and high potential benefits, RPA has been widely adopted in many industries. As of April 2015, Telefónica O2, the second-largest mobile telecommunications provider in the United Kingdom, had adopted more than 100 RPA “robots” to handle 500,000 transactions each month (Lacity et al. 2015). In addition, a business process outsourcing provider automated 14 core processes with RPA, achieving a typical 30% cost saving per process and improving service quality and accuracy. Also, in the process of updating London Premium Advice Notes (LPANs) to a central insurance market repository, an RPA robot was used to automate the most onerous steps: validating data, accessing the database, creating documents, and uploading the repository. After adoption, the processing time was only 30 minutes instead of several days (Deloitte 2017).

Even though the industry has observed the benefits of RPA, its applications in auditing practice are still unexplored. Additionally, many audit tasks are well-defined, highly repetitive, and predictable; for example, extracting exogenous information (confirmations from the electronic platform and customer reviews from social media) and matching information from multiple systems, which are multi-step tasks across multiple systems, are the ideal candidates for RPA (IRPA 2015). With the improved processing power of RPA, the scale of audit procedures can be increased and auditors will be able to focus on tasks that require professional judgment and higher order thinking skills, thereby enhancing audit quality.

CHAPTER 3: DO MANAGERS USE EXTENSION ELEMENTS STRATEGICALLY IN THE SEC'S INTERACTIVE DATA FOR FINANCIAL STATEMENTS? EVIDENCE FROM XBRL COMPLEXITY

3.1. Introduction

The "Incomplete Revelation Hypothesis" suggests that managers in poorly performing firms have more incentive to obfuscate financial information because the resulting incomplete market reaction would be the difficulty of extracting the information from firms' public disclosure (Bloomfield 2002). Consistent with this hypothesis, prior literature documents that the financial reporting of firms with strong performance tends to be straightforward (Schrand and Walther 2000) while that of firms with poor performance is more complex and difficult to understand (Li 2008). This study extends this stream of literature by investigating whether firms report financial statements strategically with eXtensible Business Reporting Language (XBRL).[1] Specifically, we examine whether firms with poor performance are more likely to issue complex XBRL filings compared with firms with good performance.

XBRL, mandated by the Securities and Exchange Commission (SEC), is a global open standard for preparing, publishing, exchanging, and consuming financial information. By tagging each fact in financial statements using pre-defined elements, XBRL makes it

---

[1] All of the Securities and Exchange Commission (SEC) filing formats in the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system are machine-readable. As the SEC states, "HTML currently is best suited for providing human-readable text" (SEC 2009, 33): machines can read firms' filings in Hypertext Markup Language (HTML) or Portable Document Format (PDF) format and deliver a human-readable presentation. HTML and PDF formats, however, do not provide meaning and context to data. XBRL is a machine-understandable format that attaches metadata (i.e., data that describes other data) to provide meaning and context to financial data. Similar to the SEC statement that "the term Interactive Data File means the machine-readable computer code that presents information in eXtensible Business Reporting Language (XBRL) electronic format "(SEC 2009, 168), this paper refers to XBRL as a machine-readable format to reflect the fact that XBRL-tagged data are understandable by machines.

possible for machines to understand what a tagged number represents based on the element used (i.e., which financial concept) and how it relates to other numbers. Hence, XBRL elements not only make financial data more accessible but also enable investors, regulators, and related parties to handle financial data more quickly, easily, and cheaply and thus improve their analysis and decision making (Dong et al. 2016; XBRL.US 2014). Under the SEC's XBRL mandate (SEC 2009), firms can define their own elements to create extension taxonomies, called extension elements, if appropriate elements do not exist in the official U.S. GAAP taxonomy. Although creating extension elements can enhance reporting flexibility (SEC 2009), the use of extension elements requires manual interpretation of tagged data, reduces reporting standardization, and hinders cross-firm comparability (Boritz and No 2008). Consistent with these concerns, both regulators and practitioners have expressed concerns about unnecessary extension elements (SEC 2010, 2011, 2014; Chasan 2013; Desmond 2014; McCafferty 2013). Researchers have also found initial evidence that managers overuse extensions (Debreceny et al. 2011) and that the overuse of extension elements may impair market efficiency (Dhole et al. 2015; Kirk et al. 2016).

Whether the overuse of extension elements is due to the firms' lack of experience with XBRL filings or their strategic judgment in choosing XBRL elements to obfuscate financial information is an interesting empirical question. On the one hand, Guragai et al. (2014) point out that together with financial statement users' inexperience in interpreting XBRL filings, managers have an opportunity to mispresent financial information with a lower chance of detection by increasing the complexity of XBRL filings. Furthermore, Hoitash and Hoitash (2017) observe that extension elements are associated with discretionary accruals and suspect that managers' overuse of extension elements reflects an attempt to obfuscate information. On the other hand, Scherr and Ditter (2017) do not find an association between the use of extension elements and earnings quality and suggest that filing complexity and the cost of compliance are the main determinants in the deviation from using standard elements. Accordingly, more research might provide insights into the debate over whether managers use extension elements strategically to introduce complexity into XBRL filings (henceforth XBRL complexity) and to obfuscate XBRL-tagged financial information.

We employ the ratio of extension elements to total elements used to tag financial facts in 10-K filings (hereafter extension ratio) as our main measure for XBRL complexity. Using a sample of mandatory XBRL 10-K filings[2] between 2009 and 2015, we find that XBRL complexity is negatively associated with firm performance. In addition, firms with less (more) persistent positive (negative) earnings have more complex XBRL filings. Such relations continue to hold after controlling for firm-specific characteristics (e.g., size, market to book value), market condition (e.g., earnings volatility), level of monitoring (e.g., analyst coverage), operational complexity (e.g., number of segments), and readability of textual financial reporting (e.g., length). Furthermore, the relation is more prominent when the firms are inherently more complex.

Overall, the findings of this study provide several contributions to the literature. First, we contribute to the line of strategic reporting literature where prior studies have examined the strategic timing, dissemination, and textual readability of disclosures to

[2] During the first two years of the SEC's mandated XBRL reporting, XBRL submissions were subject to limited liability and are deemed to have been furnished, not filed, because of the liability provisions of the Securities Act of 1933 and the Securities Exchange Act of 1934 (SEC 2009). Although our sample consists of both furnished and filed XBRL submissions, we refer to the mandated XBRL submissions as XBRL filings for simplicity.

benefit firms and managers (Li 2008; Jung et al. 2016; DeHaan et al. 2015). Different from existing research, this study explores whether firms use XBRL elements strategically to tag financial information. In particular, we provide evidence that managers tag financial information in their XBRL 10-K filings strategically by using extension elements to increase the complexity of the filings. Second, our research contributes to a recent debate in the XBRL literature. After observing an association between discretionary accruals and the use of extension elements, Hoitash and Hoitash (2017) suspect that the overuse of extension elements is a form of managerial discretion; in contrast, Scherr and Ditter (2017) identify several factors that determine the use of extension elements but do not find an association between their use and earnings quality, suggesting no evidence for their strategic use. Based on the Incomplete Revelation Hypothesis (Bloomfield 2002) and the research design of Li (2008), this study shows a relationship between XBRL complexity and firm performance. The findings enhance our understanding of the effect of managerial discretion on the choice of extension elements to tag financial information and thus provide evidence for the debate. Finally, our study offers policy implications to regulators for improving the quality of financial reporting. Since the mandatory adoption of XBRL, the SEC has issued a series of staff observation letters concerning the use of unnecessary extension elements and the consequences for the information environment. This study provides empirical evidence that managers' reporting strategies are one possible reason for unnecessary extension elements.

The remainder of the essay is organized as follows. The next section presents the research background and introduces the hypotheses. The third section describes the sample selection and research design. In the fourth section, we report the results of hypothesis tests and related sensitivity analyses. The last section summarizes the findings, implications, and limitations of the study.

3.2. Background and Hypothesis Development

3.2.1 Strategic Reporting

The strategic reporting literature has documented that managers have disclosure strategies for disseminating firm-specific information (Li 2008; Jung et al. 2016; DeHaan et al. 2015). Such a strategy includes decisions regarding the timing of the disclosure, the channels of communication, and the readability of the reports. For instance, managers strategically hide bad news through the timing of disclosure, for example by announcing adverse information during a period of low attention such as after trading hours, on Fridays, and on busy reporting days (Segal and Segal 2016; Bagnoli et al. 2005; DeHaan et al. 2015). In focusing on strategic dissemination, which is a firm's decision about whether to use a certain channel to communicate with investors, Jung et al. (2016) investigate firms' discretionary use of social media and find that firms are less likely to disseminate earnings announcements via Twitter when they release adverse information.

More related to this study, another line of literature explores how managers use the complexity of textual disclosures to mitigate a negative market reaction or to strengthen a positive market reaction. Bloomfield (2002), in proposing the Incomplete Revelation Hypothesis (IRH), argues that information is less completely reflected in market prices when extracting it from publicly available data is more costly. According to IRH, managers may have an incentive to increase the processing cost of adverse information to mitigate or delay a negative market reaction. Consistent with the arguments above, Li (2008) uses both the Fog Index and the number of words to assess financial reporting readability and shows that firms with less readable 10-Ks have lower earnings, suggesting that managers try to hide bad news from investors by increasing the complexity of documents. Following Li (2008), Lundholm et al. (2014) find that foreign firms produce more readable financial statements compared with similar U.S. firms, arguing that foreign-based firms have a greater incentive to make their public documents more readable in order to attract U.S. investors. However, this stream of literature focuses only on the processing cost of extracting relevant information from textual documents. Surprisingly, little attention has been devoted to information provided in a tagged data format (i.e., XBRL), especially considering that financial information is increasingly extracted, processed, and interpreted by machines (Moffitt and Vasarhelyi 2013).

3.2.2 XBRL Extension Taxonomy

Under the SEC's XBRL mandate, firms are allowed to create unique elements (i.e., extension elements) and provide them as an extension taxonomy if elements that meet their needs do not exist in the U.S. GAAP taxonomy. However, although the use of extension elements is likely to enhance reporting flexibility (SEC 2009), such elements require manual interpretation of the tagged data for investors and other stakeholders, and thus their use impairs reporting standardization and cross-firm comparability (Boritz and No 2008). Therefore, the SEC has raised concerns about the unnecessary use of extension taxonomies in a series of observation letters (SEC 2010, 2011, 2014). Practitioners have also paid considerable attention to the "unhealthy" ratio of extension tags and widely discussed the negative consequences for financial reporting quality (e.g., Chasan 2013; Desmond 2014; McCafferty 2013). Consistent with this concern, Debreceny et al. (2011) manually analyze the extension elements used in the primary financial statements for the filings submitted during the first year of XBRL adoption. They observe that about 40 percent of the extensions are made for available standard elements, suggesting a more appropriate choice of elements is necessary. Furthermore, prior studies document the unintended effects of extension elements. Dhole et al. (2015) show that the use of extension elements impairs the comparability of financial statements. Kirk et al. (2016) also find that the reduced comparability caused by complex XBRL filings (i.e., more use of extension elements) is likely to lead to lower analyst forecast accuracy and greater dispersion.

Since the XBRL tagging process (i.e., choosing appropriate elements and creating extension elements for financial facts) often involves significant management discretion (Kirk et al. 2016), what leads to the overuse of extension elements has been a focus in the literature. Guragai et al. (2014) suggest that opportunistic managers, realizing that financial statement users lack experience in interpreting XBRL filings, may take advantage of the complexity of XBRL filings to mispresent their disclosures. Furthermore, both Kim et al. (2013) and Hoitash and Hoitash (2017) observe that extension elements are associated with an increase in the magnitude of discretionary accruals, leading them to suspect that managers' attempts to obfuscate information may lead to the overuse of extension elements.

3.2.3 Hypothesis

Scherr and Ditter (2017) find no association between extension elements and reporting quality and conclude that extension elements are associated predominantly with filing complexity and the cost of compliance, rather than with strategic usage. However, given that the extensive use of these elements impedes reporting standardization and cross-firm comparability (Boritz and No 2008), it is conceivable that managers may strategically use extension elements to increase the complexity of the filings to mitigate or delay a negative market reaction. Therefore, whether managers use extension elements strategically is an interesting research question. Applying the same logic behind IRH (Bloomfield 2002) and prior literature on financial reporting readability (Li 2008; Loughran and McDonald 2014) to the XBRL setting, we argue that managers have an incentive to strategically increase XBRL complexity (i.e., the complexity of XBRL filings) when firms perform poorly. Accordingly, we propose the following hypothesis:

H1: Firm performance is negatively associated with XBRL complexity.

The relation between XBRL complexity and firm performance suggested by IRH can also be extended to future performance. Opportunistic managers have incentives to increase processing cost in order to reduce or delay market reactions if the good news is less persistent. In addition, to distinguish themselves from "fake" good firms, managers of firms with better future performance are motivated to increase the transparency of financial reports and thus encourage investors to fully absorb favorable information. Li (2008) finds that firms' reports are more likely to be less readable if positive earnings are transitory. Similar to our first hypothesis, we expect that managers have more incentives to strategically disclosure complex XBRL filings if the firm has a less persistent good performance or a more persistent bad performance, which leads to our second hypothesis.

H2a: XBRL filing is more complex when firms have less persistent positive earnings.

H2b: XBRL filing is more complex when firms have more persistent negative earnings.

Compared to firms with less firm-specific information, complex firms tend to have more distinctive financial facts, and their tagging processes are more likely to be involved in management discretion. Therefore, with their greater discretion, opportunistic managers are more capable of using extension elements strategically to increase XBRL complexity depending on the complexity of their business. Additionally, firm-specific events may contribute to greater information asymmetry (Aboody and Lev 2000), which reduces employees' perceptions of the likelihood of being caught and disciplined for misconduct (Werbel and Balkin 2010). The perception that investors and regulators are less likely to detect potential misconduct further increases opportunistic managers' incentives, which gives the following hypothesis.

H3: The association between firm performance and XBRL complexity is more pronounced when the firm is inherently more complex.

3.3. Research Methodology

3.3.1 Sample Selection and XBRL Complexity Measure

The SEC has mandated that public firms tag financial statements using XBRL elements for fiscal periods ending on or after June 15, 2009 over 3-year phase in period. Therefore, we first collect all mandatory XBRL 10-K filings from the SEC's Electronic Data Gathering, Analysis, and Retrieval system (EDGAR) between 2009 and 2015.[3] This study focuses on 10-K filings because prior studies suggest that investors have limited reaction to quarterly financial reporting (Li and Ramesh 2009), and thus managers have less incentive to disclose quarterly financial information strategically. The initial sample contains 34,593 firm-year observations.

---

[3] The authors thank Bill McDonald for providing the 10-K Parse files and acknowledge the SeekiNF database for providing the Gunning Fog Index for the 10-K files.

Under the SEC's XBRL mandate, firms are required to tag each financial fact (i.e., quantitative amount) in the primary financial statements and each footnote item and schedule as a single block of text (i.e., block tagging) in the first year of adoption. In subsequent years, firms are subject to detailed tagging of footnotes and schedules (e.g., tag each table and amount within each footnote item separately). To rule out a possible bias caused by the difference between block tagging and detail tagging, we eliminate 7,005 10-K filings with block tagging. Next, we remove 608 XBRL 10-K filings submitted by firms participating in the XBRL Voluntary Filing Program (SEC 2005), since voluntary XBRL adopters may have better corporate governance and fundamentally different firm characteristics from non-adopters (Boritz and Timoshenko 2014; Premuroso and Bhattacharya 2008). Such differences may cause a self-selection bias.

We then merge the sample of 10-K filings from the SEC's EDGAR system with variables from Compustat. However, the reporting periods extracted from the EDGAR system do not perfectly match those from Compustat.[4] Therefore, we employ a 7-day match window on reporting periods between these two databases. Finally, we download analyst coverage data from the I/B/E/S detailed file of individual analyst earnings estimates. After excluding 10-K filings that are not matched with Compustat data or that have missing key variables, we obtain our final sample of 16,130 filings.

Table 1 summarizes the sample selection process. Table 2 shows the sample composition: Panel A presents the year distribution while Panel B shows the industry distribution.

---

[4] Compustat often uses the end of the month as the last date of the reporting period rather the actual reporting period found in the 10-K.

Table 1 Sample Selection

Initial: 10-K filings from the SEC's EDGAR system between years 2009 to 2015	34,593
Less: 10-K filings with block tagging	(7,005)
Less: 10-K filings that are filed by VFP participants	(608)
Less: 10-K filings that are not available in Compustat	(5,506)
Less: 10-K filings that are missing key financial variables	(5,344)
Full Sample	16,130

Table 1 provides details on the sample selection process.

Table 2 Sample Composition

Panel A. Year Breakdown

Year	No. of 10-K Filings
2010	264
2011	1,289
2012	3,770
2013	3,840
2014	3,800
2015	3,167
Total	16,130

Panel B. Industry Breakdown

Industry (1-digit SIC code)		No. of Observations
0	Agriculture, Forestry, and Fishing	77
1	Mining and Construction	1200
2 & 3	Manufacturing	6,462
4	Transportation, Communications, Electric, Gas and Sanitary Service	1,181
5	Wholesale and Retail trade	1,342
6	Finance, Insurance and Real Estate	3,205
7 & 8	Services	2,555
9	Public Administration	108
Total		16,130

Table 2 presents the number of observations for each year and industry

We measure the complexity of XBRL filings (i.e., XBRL complexity) using the extension ratio (i.e., the ratio of extension elements to total elements) in each 10-K filing for several reasons. First, extension elements require manual interpretation, so their use increases the complexity of XBRL-tagged data by reducing machine readability and cross-firm comparability and increases the processing cost of XBRL-tagged data. Kirk et al. (2016) and Scherr and Ditter (2017) also claim that high extension ratios lead to increased XBRL complexity, which in turn causes difficulty for analysis. Finally, both regulators and practitioners have expressed concern that the use of unnecessary extension elements causes high extension rates, which in turn have a negative effect on financial reporting quality (Chasan 2013; Desmond 2014; McCafferty 2013; SEC 2014). Based on this discussion, we argue that the extension ratio is a reasonable proxy for measuring the complexity of XBRL filings.

3.3.2 Econometric Models

We test our first hypothesis using model (1) to examine the relationship between firm performance (*Earnings*) and XBRL 10-K filing complexity (*Exratio*) (Li 2008). *Earnings* captures the firm's scaled operating income, and *Exratio* is the ratio of extension elements to total elements in the XBRL 10-K filing.

$$
\begin{aligned}
Exratio_t = {} & \beta_1 + \beta_2 Earnings_t + \beta_3 LnMVE_t + \beta_4 MTB_t + \beta_5 EarnVol_t + \beta_6 NBSeg_t \\
& + \beta_7 NGSeg_t + \beta_8 Foreign_t + \beta_9 Merger_t + \beta_{10} SI_t + \beta_{11} NAnalyst_t \\
& + \beta_{12} XBRLage_t + \beta_{13} Length_t + \beta_{14} Grosize_t \\
& + \sum \beta_i \, Industry\ and\ year\ Indicators + \varepsilon_t \qquad (1)
\end{aligned}
$$

We follow the corporate disclosure and XBRL literature to control for determinants that may affect the use of extension elements (e.g., Boritz and Timoshenko 2014; Du et al.

2013; Francis et al. 2008; Kirk et al. 2016). For firm characteristics, we include firm size (*LnMVE*), market-to-book ratio (*BTM*), earnings volatility (*EarnVol*), acquisitions (*Merger*), and special items (*SI*). Suggested by the SEC's observation letters that XBRL filing experience may affect a firm's tagging behavior (SEC 2014), we control for a firm's experience in XBRL filing (*XBRLage*). In addition, prior literature argues that external monitoring may affect disclosure quality (Ball et al. 2012). Therefore, we use the number of analysts following (*NAnalyst*) to capture the level of external monitoring. Kirk et al. (2016) suggest that the process of choosing standard elements or extension elements for tagging financial facts often involves significant management discretion. Hence, managers' discretionary use of elements is likely to be affected by a firm's inherent complexity (Hoitash and Hoitash 2017). In other words, if a firm's business is inherently more complex, the financial statements of the firm tend to have more distinctive financial facts, and thus firm managers are more likely to use extension elements to tag the facts. To that end, we control two categories of firm complexity. First, to capture the complexity of operation we use the number of business segments (*NBSeg*), the number of geography segments (*NGSeg*), and the presence of foreign transactions (*Foreign*) (Boritz and Timoshenko 2014). Second, because firms' inherent complexity will be reflected in their financial statements, we follow Li (2008) and Loughran and McDonald (2014) to measure linguistic complexity using the length of the 10-K filing (*Length*) and the gross 10-K filing size (*Grosize*).[5] Appendix A contains the definition of each variable.

Although we have included the control variables above in the model (1), we cannot

[5] Loughran and McDonald (2014) argue that the Fog Index, a widely used readability measure, is not appropriate in business writing and suggest using the gross 10-K filing size (*Grosize*). We find similar results if we include the Fog Index (*FogIndex*) as a control variable. See details in the robustness tests section.

rule out unobservable factors affecting managers' discretionary use of extension elements. Bias caused by unobservable characteristics can be eliminated using a change model specification if the unobservable differences remain invariant during the period of study (Lennox et al. 2011). Following Li (2008), we adopt a year-to-year change research design to further confirm the relation between firm performance and XBRL complexity. Model (2) shows our year-to-year change model. All variables in the model (2) are the first-order difference of variables from the model (1) except for *XBRLage*.[6]

$$\begin{aligned} D.Exratio_t = {} & \beta_1 + \beta_2 D.Earnings_t + \beta_3 D.LnMVE_t + \beta_4 D.MTB_t + \beta_5 D.EarnVol_t \\ & + \beta_6 D.NBSeg_t + \beta_7 D.NGSeg_t + \beta_8 D.Foreign_t + \beta_9 D.Merger_t \\ & + \beta_{10} D.SI_t + \beta_{11} D.NAnalyst_t + \beta_{12} XBRLage_t + \beta_{13} D.Length_t \\ & + \beta_{14} D.Grosize_t + \sum \beta_i \, Industry\ and\ year\ Indicators + \varepsilon_t \quad (2) \end{aligned}$$

In addition, we employ model (3) to examine the effect of XBRL complexity on earnings persistence. The dependent variable is the earnings in year t+1 ($Earnings_{t+1}$), and the independent variables are current earnings ($Earnings_t$), XBRL complexity ($Exratio_t$), and their interaction ($Earnings_t * Exratio_t$). The interaction term is the variable of interest, as it reflects the changes in earnings persistence as XBRL complexity changes. The control variables are the same as those in the model (1). In addition, the absolute amount of accruals ($AbsAcc_t$) and a dividend payment dummy (Div_t) are also included as control variables since the literature has documented their relationship with earnings persistence (Skinner and Soltes 2011; Sloan 1996). Also, model (3) includes the interactions between current earnings ($Earnings_t$) and all control variables.

[6] *XBRLage* has the same year-to-year variance across all firms. Therefore, only *XBRLage* is included in the Model (2).

$$
\begin{aligned}
Earnings_{t+1} = {} & \beta_1 + \beta_2 Earnings_t + \beta_3 Exratio_t + \beta_4 Earnings_t * Exratio_t \\
& + \beta_5 LnMVE_t + \beta_6 MTB_t + \beta_7 EarnVol_t + \beta_8 NBSeg_t + \beta_9 NGSeg_t \\
& + \beta_{10} Foreign_t + \beta_{11} Merger_t + \beta_{12} SI_t + \beta_{13} NAnalyst_t \\
& + \beta_{14} XBRLage_t + \beta_{15} Length_t + \beta_{16} Grosize_t + \beta_{17} AbsAcc_t \\
& + \beta_{18} Div_t + \sum \beta_i \, Earnings_t * Controls_t \\
& + \sum \beta_j \, Industry\ and\ year\ Indicators + \varepsilon_t \qquad (3)
\end{aligned}
$$

To test our third hypothesis, we use operational complexity and linguistic complexity to measure the firms' inherent complexity (Hoitash and Hoitash 2017). A principal component analysis (PCA) is employed to generate an operating complexity factor and a linguistic complexity factor (Feng et al. 2009) because each complexity-related variable might reflect a different dimension of operating or linguistic complexity. Specifically, we aggregate the number of business segments (*NBSeg*), the number of geographic segments (*NGSeg*), and the presence of foreign business (*Foreign*) into the operating complexity factor (*OpeComplex*). We combine the length of the 10-K filing (Length) and the gross 10-K filing size (*Grosize*) to form the linguistic complexity factor (*LinComplex*). Then, in models (4) and (5), we include each complexity measure and its interactions with *Earnings* to examine the incremental effect of firm performance on XBRL complexity as firms' complexity changes. The dependent variable and control variables are those used in the model (1).

$$Exratio_t = \beta_1 + \beta_2 Earnings_t + \beta_3 LnMVE_t + \beta_4 MTB_t + \beta_5 EarnVol_t + \beta_6 Merger_t + \beta_7 SI_t + \beta_8 NAnalyst_t + \beta_9 XBRLage_t + \beta_{10} Length_t + \beta_{11} Grosize_t + \beta_{12} OpeComplex_t + \beta_{13} Earnings * OpeComplex_t + \sum \beta_i \, Industry\ and\ year\ Indicators + \varepsilon_t \quad (4)$$

$$Exratio_t = \beta_1 + \beta_2 Earnings_t + \beta_3 LnMVE_t + \beta_4 MTB_t + \beta_5 EarnVol_t + \beta_6 NBSeg_t + \beta_7 NGSeg_t + \beta_8 Foreign_t + \beta_9 Merger_t + \beta_{10} SI_t + \beta_{11} NAnalyst_t + \beta_{12} XBRLage_t + \beta_{13} LinComplex_t + \beta_{14} Earnings * LinComplex_t + \sum \beta_i \, Industry\ and\ year\ Indicators + \varepsilon_t \quad (5)$$

3.4 Results

3.4.1 Descriptive statistics

Panel A of Table 3 displays the summary statistics of key variables. XBRL complexity (*Exratio*) exhibits variance with a ratio of 0.14 at the 25th percentile and 0.25 at the 75th percentile. The average natural logarithm of total words (*Length*) is 11, which is similar to the findings in Hoitash and Hoitash (2017). The average 10-K filing size (*Grosize*) in our sample is 14.8 megabytes, which is significantly larger than the mean value of 10-K filing size in Loughran and McDonald (2014). However, a major increase in 10-K filing size after 2009 is expected because of the SEC’s XBRL mandate. Under the SEC mandate, firms are required to provide the XBRL format of financial statements in addition to the previous filing formats such as HTML and PDF.

Panel B of Table 3 shows the Pearson correlation matrix of our key variables. XBRL complexity (*Exratio*) is negatively related to firm performance (*Earnings*), providing initial evidence to support our hypotheses. Consistent with our expectation, XBRL complexity is negatively associated with firms' XBRL filing experience (*XBRLage*) and positively associated with the linguistic complexity measures (*Grosize* and *Length*). However, the association between the operational complexity proxies and XBRL complexity is mixed.

Table 3 Summary Statistics

Panel A. Descriptive Statistics

Variable	N (16,130)				
	Mean	Median	25%	75%	Std. Dev.
Exratio	0.20	0.19	0.14	0.25	0.09
Earnings	-0.20	0.04	-0.01	0.10	1.15
LnMVE	6.18	6.41	4.41	7.92	2.41
MTB	3.56	1.48	1.06	2.41	9.56
EarnVol	1.13	0.05	0.02	0.11	6.43
NBSeg	1.70	1	1	2	1.14
NGSeg	1.76	1	1	2	1.63
Foreign	0.28	0	0	1	0.45
Merger	0.13	0	0	0	0.34
SI	-0.02	0.00	-0.01	0.00	0.08
NAnalyst	1.61	1.79	0.00	2.64	1.23
XBRLage	3.50	3	2	4	1.35
Grosize (Mb)	16.49	16.51	16.09	16.94	0.65
Length	11.00	11.00	10.71	11.30	0.54

Panel B. Pearson Correlation Matrix

Variable	*Exratio*	*Earning*	*LnMVE*	*MTB*	*EarnVol*	*NBSeg*	*NGSeg*	*Foreign*	*Merger*	*SI*	*NAnalys*	*XBRLag*	*Grosize*	*Length*
Exratio	1													
Earnings	-0.071	1												
LnMVE	0.119	0.362	1											
MTB	0.006	-0.823	-0.233	1										
EarnVol	0.052	-0.437	-0.229	0.382	1									
NBSeg	0.044	0.132	0.262	-0.117	-0.085	1								
NGSeg	-0.109	0.086	0.106	-0.062	-0.067	-0.097	1							
Foreign	-0.080	0.079	0.176	-0.059	-0.040	0.096	0.312	1						
Merger	-0.031	0.075	0.142	-0.060	-0.032	0.112	0.043	0.083	1					
SI	-0.010	0.197	0.127	-0.124	-0.101	0.022	-0.002	-0.020	-0.004	1				
NAnalyst	0.046	0.274	0.786	-0.193	-0.187	0.180	0.146	0.169	0.123	0.082	1			
XBRLage	-0.123	0.042	0.237	-0.037	-0.045	0.071	0.001	0.040	0.015	0.008	0.183	1		
Grosize	0.267	0.324	0.582	-0.339	-0.221	0.244	0.023	0.107	0.082	0.073	0.418	0.129	1	
Length	0.343	0.178	0.482	-0.209	-0.138	0.167	0.001	0.082	0.094	0.012	0.381	0.114	0.616	1

Table 3 presents descriptive statistics and correlation matrix for the dependent, independent, and control variables included in our analysis. See Appendix A for variable definitions.

3.4.2 Main Findings

Table 4 reports the regression results of the test of the first hypothesis. The variable of interest is *Earnings*, which measures firm performance. The coefficient on *Earnings* in the Level Specification column is negative and significant (-0.013; $p < 0.01$), indicating that firms' XBRL filings contain more extension elements when the firms are performing poorly. The coefficients on the control variables are mainly consistent with our expectations. The coefficient on *LnMVE* (natural logarithm of market value) is negative and significant, which is consistent with the SEC's observation that smaller filers have higher extension ratios (SEC 2014). The linguistic complexity measures of the 10-K filings (*Grosize* and *Length*) are positivity related to XBRL complexity (*Exratio*), suggesting that the complexity of XBRL filings is affected by the complexity of the financial statements.

Table 4 Regression of Firm Performance on XBRL Complexity

Level Specification		Change Specification	
Independent Variables	Dependent Variable	Independent Variables	Dependent Variable
	$Exratio_t$		*$D.Exratio_t$*
Intercept	-0.448***	*Intercept*	-0.058***
	(-9.22)		(-9.68)
$Earnings_t$	-0.013***	*$D.Earnings_t$*	-0.003***
	(-6.27)		(-2.61)
$LnMVE_t$	-0.002**	*$D.LnMVE_t$*	0.000
	(-2.54)		(0.46)
MTB_t	-0.001***	*$D.MTB_t$*	-0.000**
	(-2.59)		(-2.41)
$EarnVol_t$	0.001**	*$D.EarnVol_t$*	0.000
	(2.38)		(1.06)
$NBSeg_t$	-0.001	*$D.NBSeg_t$*	0.000
	(-1.04)		(0.44)
$NGSeg_t$	-0.002***	*$D.NGSeg_t$*	0.000
	(-2.98)		(0.17)
$Foreign_t$	-0.008***	*$D.Foreign_t$*	0.001
	(-3.36)		(0.52)
$Merger_t$	-0.007***	*$D.Merger_t$*	-0.001
	(-3.68)		(-1.19)
SI_t	-0.006	*$D.SI_t$*	0.002
	(-0.57)		(0.22)
$NAnalyst_t$	-0.003**	*$D.NAnalyst_t$*	-0.003***
	(-2.55)		(-3.26)
$XBRLage_t$	0.003**	*$XBRLage_t$*	-0.001***
	(2.13)		(-2.76)
$Grosize_t$	0.023***	*$D.Grosize_t$*	0.011***
	(8.58)		(5.82)
$Length_t$	0.039***	*$D.Length_t$*	0.004***
	(14.39)		(5.40)
Year and Industry fixed	Included	Year and Industry fixed	Included
N	16,130	N	11,102
Adj-R^2	29.9%	Adj-R^2	6.2%

Table 4 reports the results of the level and change regression models (H1). Reported t-statistics are estimated with the clustered standard error by firms. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

Considering the endogeneity issue caused by unobservable factors, we also test our first hypothesis using model (2), a change model specification. As shown in the Change Specification column, the coefficient of year-to-year change in firm performance (*D.Earnings*) is negative and significant (-0.003; $p < 0.01$). This result rules out the possible effect of unobservable factors and further confirms the negative relation between firm performance and XBRL complexity.

One alternative explanation for our findings is that firms with bad performance have more unique events than routine transactions to disclose. Since standard XBRL elements are not designed to tag ad-hoc events, those firms have to use more extension elements. Our argument for this explanation is that if a firm has more unique events to disclose because of bad performance, the linguistic complexity of its financial statements should be able to capture the variance. As the variables measuring the linguistic complexity of the firms' financial statements are included as control variables in the models above, this explanation is being addressed. Collectively, the evidence in both the level and change specifications supports our first hypothesis that firms with worse performance have more complex XBRL filings. That is, our results suggest that managers tend to use extension elements strategically to obfuscate XBRL-tagged financial information.

Table 5 shows the regression results for the tests of the second hypothesis. Our variable of interest is *Earnings*Exratio*, which captures the changes in earnings persistence as XBRL complexity changes. We use all firm-years with positive (negative) earnings to test the persistence of good (bad) news. As shown in Table 5, the coefficient of

*Earnings*Exratio* in the Positive Earnings column is negative and significant (-0.774; $p < 0.05$), indicating that earnings in year t+1 are less associated with current earnings when XBRL filing is complex. The result supports Hypothesis 2a, indicating that XBRL filing is more complex when good news is less persistent. The coefficient of *Earnings*Exratio* in the Negative Earnings column is positive but only marginally significant ($p = 0.11$), partially supporting Hypothesis 2b. Untabulated results reveal similar findings concerning the association between XBRL complexity and earnings persistence in year t+2. Taken together, the results in Table 5 suggest that as positive (negative) earnings are less (more) persistent, XBRL complexity increases.

Table 5 Regression of XBRL Complexity on Earnings Persistence

	Positive Earnings	Negative Earnings
Independent Variables	Dependent Variable	
	$Earnings_{t+1}$	
Intercept	0.201**	-0.17
	(2.09)	(-0.15)
$Earnings_t$	0.325	2.125**
	(0.36)	(2.16)
$Exratio_t$	0.04	0.286
	(1.51)	(1.05)
$Earnings_t$ * $Exratio_t$	-0.774**	0.417
	(-2.19)	(1.62)
$LnMVE_t$	0.004	0.032
	(1.28)	(1.51)
MTB_t	0.000	-0.034***
	(0.16)	(-3.92)
$EarnVol_t$	-0.000	-0.007*
	(-0.28)	(-1.85)
$NBSeg_t$	0.004**	0.000
	(2.12)	(0.01)
$NGSeg_t$	0.002	0.015
	(1.14)	(1.02)
$Foreign_t$	0.008*	0.071*
	(1.68)	(1.71)
$Merger_t$	0.013***	0.048
	(3.35)	(1.29)
SI_t	-0.165	0.097
	(-1.16)	(0.28)
$NAnalyst_t$	-0.003	-0.022
	(-0.76)	(-0.79)
$XBRLage_t$	-0.003	-0.012
	(-1.08)	(-0.24)
$Grosize_t$	-0.006	0.057
	(-1.12)	(0.93)
	-0.011**	-0.123**

$Length_t$	(-2.39)	(-2.25)
$AbsAcc_t$	-0.001	-0.299
	(-0.02)	(-1.45)
Div_t	-0.007	0.059
	(-1.47)	(0.72)
Interactions with control variables	Included	Included
Year and Industry fixed effects	Included	Included
N	8,311	2,791
Adj-R^2	35.6%	60.8%

Table 5 reports the results of the association of earning persistence and XBRL complexity (H2). Reported t-statistics are estimated with the clustered standard error by firms. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

Table 6 shows the regression results of the tests of the third hypothesis. The variables of interest are the interaction between the complexity measures and firm performance (*OpeComplex*Earnings* and *LinComplex*Earnings*). The coefficient of *OpeComplex*Earnings* in the Effect of Operational Complexity column is negative and significant (-0.010; $p < 0.05$). The Effect of Linguistic Complexity column provides the regression results for the effect of linguistic complexity, which is also negative and significant (-0.004; $p < 0.01$). These results support H3, revealing that the negative association between firm performance and XBRL complexity is more pronounced when the firm is more operationally complex, or the firm's financial statements are more textually complex.

Table 6 Regression of Accounting Complexity on XBRL Complexity

Effect of Operational Complexity		Effect of Linguistic Complexity	
Independent Variables	Dependent variable	Independent Variables	Dependent
	Exratio$_t$		*Exratio$_t$*
Intercept	-0.447*** (-9.25)	*Intercept*	0.357*** (12.91)
Earnings$_t$	-0.015*** (-6.81)	*Earnings$_t$*	-0.017*** (-6.82)
LnMVE$_t$	-0.002** (-2.49)	*LnMVE$_t$*	-0.002*** (-2.70)
MTB$_t$	-0.001** (-2.50)	*MTB$_t$*	-0.000* (-1.87)
EarnVol$_t$	0.001** (2.42)	*EarnVol$_t$*	0.001*** (2.59)
Merger$_t$	-0.007*** (-3.75)	*NBSeg$_t$*	-0.001 (-1.22)
SI$_t$	-0.005 (-0.51)	*NGSeg$_t$*	-0.002*** (-2.99)
NAnalyst$_t$	-0.003** (-2.50)	*Foreign$_t$*	-0.009*** (-3.54)
XBRLage$_t$	0.004** (2.16)	*Merger$_t$*	-0.007*** (-3.64)
Grosize$_t$	0.023*** (8.52)	*SI$_t$*	-0.003 (-0.30)
Length$_t$	0.038*** (14.26)	*NAnalyst$_t$*	-0.003** (-2.52)
OpeComplex$_t$	-0.011*** (-5.11)	*XBRLage$_t$*	0.003** (2.05)
*OpeComplex$_t$ * Earnings$_t$*	-0.010** (-2.44)	*LinComplex$_t$*	0.043*** (21.56)
		*LinComplex$_t$ * Earnings$_t$*	-0.004*** (-2.97)
Year and Industry fixed	Included	Year and Industry fixed	Included
N	16,130	N	16,130
Adj-R^2	30.0%	Adj-R^2	30.0%

Table 6 reports the results of the incremental effects of operational complexity and linguistic complexity (H3). Reported t-statistics are estimated with the clustered standard error by firms. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

3.4.3 Robustness Tests

Several robustness tests were conducted. Hoitash and Hoitash (2017) have proposed using total elements as a new accounting reporting complexity (ARC) measure and claim that ARC is more persistent than prior complexity measures. Their findings suggest that our complexity measure (i.e., the extension ratio) may capture not only XBRL complexity but also overall accounting complexity. To address the potential confounding effect resulting from our measure, we include the ARC measure suggested by Hoitash and Hoitash (2017) to control for the overall accounting complexity in all our models. We observe similar regression results, which further supports our main findings and the validity of our XBRL complexity measure.

Loughran and McDonald (2014) suggest using the gross size of the 10-K files (*Grosize*) to measure the readability of the 10-K and argue that the Fog Index (*FogIndex*) is not an appropriate readability measure in business writing. However, in the accounting literature, the Fog Index is still a widely used proxy for financial statement readability (e.g., Lawrence 2013; Lehavy et al. 2011; Lundholm et al. 2014). Therefore, we conduct two sets of sensitivity tests to eliminate the potential bias caused by different readability measures. First, we add the Fog Index (*FogIndex*) as another linguistic complexity measure in our models and generate a linguistic complexity factor (*LinComplex*) using *FogIndex*, *Grosize*, and *Length*. Second, we replace *Grosize* with *FogIndex* in our models and aggregate only *FogIndex* and *Length* into *LinComplex*. The regression results are similar to our main findings.

Finally, the SEC suggests that the use of unnecessary extension elements may be attributable to different reporting tools (SEC 2014). To address the alternative explanation that our findings are driven by reporting tools (e.g., Fujitsu's XWand and Workiva's Wdesk), we include dummy variables in all models to capture which tools are employed by firms to prepare their XBRL filings. We find qualitatively similar results, which further confirms our main findings.

3.5. Conclusion

Increasing amounts of financial information are being extracted, processed, and interpreted automatically by machines. XBRL was introduced to provide meaning and context to financial data by tagging each financial fact in financial statements using pre-defined elements. Its use enables machines to understand the meaning and context of the tagged data: what the number represents and how it relates to other numbers. Prior literature shows the overuse of extension elements and attempts to address whether managers overuse extension elements strategically. However, existing studies disagree: Hoitash and Hoitash (2017) suspect that managers' attempts to obfuscate information leads to the overuse of extension elements whereas Scherr and Ditter (2017) argue that the use of extension elements is not related to managerial discretion. In addition, most studies in the strategic reporting literature center mainly on the timing, dissemination, and narratives of disclosures, not on the tagged data format of financial information.

This study focuses on a current debate in the literature and investigates whether managers use XBRL elements strategically to increase the complexity of the tagged data format of financial statements (i.e., XBRL filings). The SEC's XBRL mandate, which requires firms to tag each financial fact in financial statements using either a standard element or an extension element, offers a unique setting for our research questions and allows us to examine managers' discretionary use of extension elements without a self-selection bias. Using the ratio of extension elements to total elements in an XBRL 10-K filing as the measure of XBRL complexity, we find evidence that firms' XBRL filings are more complex when the firms are performing poorly and when their good news (bad news) is less (more) persistent. We further show that the effect is more pronounced when firms are inherently more complex. Our results suggest that managers tend to use extension elements strategically to introduce XBRL complexity and obfuscate XBRL-tagged financial information.

Our study makes contributions to the accounting literature in three ways. First, we extend the strategic reporting literature by demonstrating that firms use extension elements strategically to increase the complexity of XBRL-tagged data. Second, this paper contributes to a recent debate in XBRL studies about whether managers use extension elements strategically. We provide initial evidence that they do. Finally, both regulators and practitioners have expressed concerns about the overuse of extension elements and its consequences (e.g., less readability, less comparability, and lower reporting quality). Our study offers regulators the useful insight that the use of unnecessary extensions may be affected by firms' reporting strategies.

This research is subject to the following caveats. First, whether firms prepare their XBRL filings in-house or outsource them because of the lack of relevant information is not controlled in our research models. Nevertheless, this paper includes control variables that capture the software used in the XBRL-tagging process and, at least partially, eliminates the effect of the preparation process on XBRL complexity. Since the tagging process may differ depending on the XBRL implementation approach (i.e., in-house vs. outsourcing), another avenue for future research is to examine whether the XBRL implementation approach affects managers' strategic use of extension elements. Second, this research focuses only on the extension ratio rather than on managers' discretionary use of elements for specific financial facts. It would be interesting for future research to explore how and based on what criteria managers choose official elements or extension elements for different financial facts.

CHAPTER 4: HOW DOES INFORMATION PROCESSING EFFICIENCY RELATE TO INVESTMENT EFFICIENCY? EVIDENCE FROM XBRL ADOPTION

4.1 Introduction

A wealth of literature documents that agency conflicts and information asymmetry between managers and investors lead firms to miss optimal levels of investment (e.g., Hubbard 1997; Stein 2003). Recently, a stream of literature argues that high financial report quality can improve capital investment efficiency because it reduces information asymmetry and attenuates the motivations of managers' myopic decisions (e.g., Bushman and Smith 2001; Healy and Palepu 2001; Lambert et al. 2007; Biddle and Hilary 2006; Biddle et al. 2009; Chen et al. 2011; Jung et al. 2014; Gomariz and Ballesta 2014). Following the same notion, other studies extend this stream of research by investigating the effect of the disclosure of material weaknesses in internal control over financial reporting (Cheng et al. 2013) and the mandated adoption of International Financial Reporting Standards (IFRS) on investment efficiency (Biddle et al. 2011; Lenger et al. 2011; Hou et al. 2016). However, these studies focus on the quality of financial reporting content but overlook the efficiency of processing information. Successful communication between managers and investors through financial reports is determined by a combination of the quality of the information content and the information processing cost. This paper suggests that the cost of processing financial information is another important factor affecting the degree of information asymmetry and managers' behavior on investment decisions. Specifically, we use the adoption of the eXtensible Business Reporting Language (XBRL) as an exogenous shock to examine whether the information processing cost affects corporate investment efficiency.

As an interactive disclosure system, XBRL is expected to significantly reduce information-processing cost and improve the accessibility and usefulness of the current EDGAR reporting system. It utilizes a list of pre-defined elements to identify each piece of data, which can then be accessed and interpreted by XBRL-compatible programs (SEC 2009). Specifically, XBRL reduces the cost of processing information by minimizing manual intervention, improving the reporting standardization, and expediting the comparison between industry peers. As suggested by prior literature (Blankespoor et al. 2014; Kim et al. 2012; Liu et al. 2014; Dong et al. 2016), the adoption of XBRL is an exogenous shock that significantly changes information processing cost and provides a natural research opportunity, free of endogeneity, to investigate the effect of information processing cost on investment efficiency.

We first posit that, after the adoption of XBRL, the reduced information processing cost leads to more efficient investments by mitigating the degree of information asymmetry and enhancing the (perceived) monitoring of managerial investment behaviors. With the benefits of XBRL on searching, extracting, and comparing firm-specific information, small investors are able to conduct basic analyses and execute their power as external "watchdogs". For sophisticated investors, such as institutional investors and analysts, the XBRL adoption optimizes their resource allocations by shifting more resources from tedious information collection to analyses (Liu et al. 2014). Thus, both small and large investors may be better able to monitor management performance, and detect managerial opportunism and possible misalignments in an effective and timely manner in the post-XBRL era (Ferreira and Matos 2008; Chang et al. 2009). With the fear of involuntary replacement and the desire to receive positive recognition from the market (Holmstrom 1982; Ali and Zhang 2015), managers may hesitate to maximize self-interests and mitigate sub-optimal investment decisions.

To test our hypothesis, we retrieve firms' XBRL filings from EDGAR's database of Interactive Data Filings and Really Simple Syndication (RSS) feeds, and identify the first XBRL filing date as the adoption date. Exploiting the merit of XBRL adoption, we regress the dependent variable representing abnormal investments (XINV), which is derived from the model suggested by McNichols and Stubben (2008), on a dummy variable (*Post*) that indicates the pre and post period of XBRL adoption. As expected, we find that the adoption of XBRL reduces the level of abnormal investments. Translated into the economic magnitude, this effect represents a 15.8% marginal reduction in inefficient investments.

In addition, we perform several analyses considering potential moderating factors that may magnify or mitigate the benefits of XBRL on the improvement in investment efficiency, including external monitoring, environmental uncertainty, and the reporting readability. First, we expect that the benefits of XBRL mandate should be more evident for firms with weaker external monitoring. If the existing governance is strong enough to provide considerable monitoring of managerial behaviors, the substitution effect will reduce the incremental gains from the adoption of XBRL on investment efficiency. Our empirical results show that the benefits driven by the XBRL adoption on investment efficiency are mitigated by the existing superior external monitoring (proxied by institutional holdings or analyst coverage) (Kim et al. 2012; Khurana and Moser 2012).

Second, we expect that information environment uncertainty also plays an important role in moderating the association between the adoption of XBRL and investment efficiency. Uncertainty in the information environment makes it more difficult for investors to assess firm-specific information and to detect opportunistic behaviors (Lim et al. 2008). Therefore, the effect of enhanced accessibility may eventually result in a higher improvement of investment efficiency. We find evidence that the benefits of XBRL adoption on investment efficiency are magnified by a firm's uncertain environment (proxied by the number of business segments and the consensus analyst forecast errors), suggesting that the improvement in investment efficiency after the XBRL mandate is more pronounced for firms with higher environmental uncertainty.

Moreover, we examine how the readability of financial reporting affect the benefits of XBRL adoption on investment efficiency. As suggested in the prior literature (Hoitash and Hoitash 2017), linguistic complexity increases external users' difficulty in consuming the reports and leads to lower information quality. Since the quality of communication via financial reports is determined by both the quality of information and processing cost, we expect that the adoption of XBRL will be more able to improve information processing efficiency and reduce information asymmetry for a company with more readable reports. Our empirical results show that the readability of financial reporting (proxied by both the Fog index and the length of reporting (Li 2008)) magnifies the benefits of XBRL adoption on investment efficiency.

We further confirm our results by conducting several additional tests. Specifically, these analyses include the "difference-in-difference" (DID) robustness check, year-to-year change design model, the dynamic effects test, non-capital investment and the effect on over- and under-investment. We find that our empirical results are robust in the more rigorous DID empirical setting and change design. To test the dynamic effects of XBRL on investment efficiency, we replace the dummy variable (*Post*) to three event year dummies (*Post1*, *Post2*, and *Post*3) and find a persistent and increasing pattern of the XBRL effect, supporting our expectation that both investors and firms face a learning curve in understanding XBRL and the effect of XBRL adoption on investment efficiency enhances as time goes. While our main hypotheses examine the effect of XBRL adoption on firm-specific capital investment efficiency, we also investigate such effects on non-capital investment (e.g., R&D and acquisitions). We also find a consistent result that the level of abnormal non-capital investments is attenuated after the adoption of XBRL. Thus, the results of robustness tests are consistent with our main findings and show that the reduced information processing cost from the XBRL mandate leads to enhanced investment efficiency for both capital and non-capital investments. Finally, we split the abnormal investment into over-investment (*XINV*>=0) and under-investment (*XINV*<0) groups, and find that the XBRL mandate is more likely to curb managerial opportunistic over-investments.

This study contributes to the existing literature in the following ways. First, this paper extends the growing literature on the benefits of XBRL adoption. Prior studies focus on the benefits of XBRL adoption mainly from the perspective of information users. For example, Kim et al. (2012) find that the reduction in investors' information processing cost after XBRL adoption mitigates information risk. Dong et al. (2016) find that XBRL adoption improves the ability of investors to incorporate firm-specific information into stock prices. Liu et al. (2014) show that the adoption of XBRL also increases analyst coverage and improves analyst forecast accuracy. Instead, our study extends the XBRL literature by providing empirical evidence on the effects of XBRL adoption from the perspective of managerial behaviors. Specifically, our findings suggest that the diminished information processing cost and the enhanced (perceived) monitoring after XBRL adoption improve manager's investment decisions.

Second, we add values to the existing accounting literature that investigates the association between financial reporting quality and firms' investment efficiency, following Biddle and Hilary (2006) and Biddle et al. (2009). Instead of exploiting the quality of financial information (Chen et al. 2011; Cheng et al. 2013), we suggest that the efficiency of processing information also plays an important role in assessing the success of communication between companies and investors. The reduced information processing cost narrows the information gap between managers and investors, and disciplines managers to make better investment decisions. To our best knowledge, our paper is the first to investigate how the cost of processing the content in financial reports, rather than the quality of the information itself, affects firm-specific investment efficiency.

Third, this study utilizes the adoption of XBRL as an exogenous shock to measure the unobservable factor: information processing cost. Such setting provides us with a natural empirical design that is free of endogeneity and a clear prediction about the effect of information processing cost on investment efficiency.

The remainder of the chapter proceeds as follows: Section 2 reviews prior literature and develops hypotheses. The research design is discussed in section 3. Section 4 presents and interprets empirical results. Additional tests are documented in Section 5, and conclusions are summarized in section 6.

4.2. Literature Review and Hypothesis Development

Prior literature suggests that deviations from firm-specific optimal investment levels can be driven by information asymmetry between managers and investors. Specifically, two primary problems are identified in prior research: moral hazards and adverse selection. A moral hazard arises when information asymmetry makes monitoring mechanisms costly, allowing managers to maximize their own self-interest by making decisions that may not be optimal for shareholders (Jensen and Meckling 1976). In terms of investment, the consequence can either be over-investment in the context of managerial empire building (Jensen 1986; Richardson 2006; Hope and Thomas 2008) or under-investment due to the manager's preference to be "effort-adverse" and lead a "quiet life" (Bertrand and Mullainathan, 2003). Information asymmetry between managers and outside investors can also create adverse selection. In this situation, better-informed managers may over-invest if they use their information advantage to time the market, issue overpriced securities, and earn excess funds (Baker, Stein, and Wurgler, 2003; Hovakimian and Hovakimian, 2009).

From the agency theory perspective, there are several control mechanisms, such as corporate governance (Bushman and Smith 2001) and financial disclosure (Healy and Palepu 2001; Lambert et al. 2007; Hope and Thomas 2008), to reduce information asymmetry and facilitate better supervision of managers' potential opportunistic behaviors. Many empirical studies shed light on how financial report quality can improve investment efficiency by mitigating information asymmetry (e.g., Biddle and Hilary 2006; Biddle et al. 2009; Chen et al. 2011; Cheng et al. 2013). However, the discussions of financial reporting quality in these studies focus on the quality of content but overlook the cost of processing such content. Successful communication between managers and investors through financial reports is determined by a combination of the information quality and processing cost. This paper suggests that the cost of processing financial information is another factor that affects the degree of information asymmetry and disciplines managers to make better investment decisions. This study treats the adoption of XBRL as an exogenous shock in order to measure the unobservable factor, information processing cost, and to examine whether the efficiency of processing financial information has an effect on managers' investment behaviors.

In the 1990s, the SEC initiated the EDGAR disclosure system to make financial statements more accessible to investors through the widespread use of the Internet. To enhance the current EDGAR reporting system with a more interactive disclosure format, the SEC mandated the use of XBRL in 2009, expecting to reduce the cost of information processing and improve the accessibility and usefulness of financial statements (SEC 2009). XBRL is an open standard system that seeks to improve the preparing, publishing, communicating, and processing of financial data. An XBRL filing uses a list of pre-defined elements to identify each piece of data in the report, which then allows the data to be accessed and analyzed by XBRL-compatible programs. By minimizing manual processes, machine-readable XBRL filings reduce the cost of accessing, extracting, and interpreting financial information (Kim et al. 2012). In addition, the summary of pre-defined list of tags is called standard taxanomy, which promotes the standardization of reporting systems and facilitates the cross-company comparability of fundamental capital ratios (Boritz and No 2009). Based on the information environment changes brought by the adoption of XBRL, a growing number of empirical studies have extensively documented the benefits of the

XBRL adoption in terms of the enhanced transparency, reduced information asymmetry, and improved accessibility of financial data (e.g., Blankespoor et al. 2014; Kim et al. 2012; Liu et al. 2014).

The adoption of XBRL enhances investors' information-processing capacity. The mandated XBRL filings give small investors[7] more accessible financial information in a user-friendly and less costly search-facilitating information environment (SEC 2009). Compared to sophisticated investors, small investors generally have fewer resources and limited ability to process information (Blankespoor et al. 2014), which constrains their decision-making ability. Since the adoption of XBRL makes it simpler to search, extract, and compare firm-specific information, it should be much easier for small investors to conduct basic analysis, such as comparing financial ratios among competitors, evaluating suggestions from other information channels, and generating their own opinions based on the firm-specific information (XBRL.US 2009). Consistent with this view, Hodge et al. (2004) use behavioral experiments to show that non-professional users are likely to benefit from search-facilitating technologies like XBRL in analyzing financial reports. In addition, the adoption of XBRL brings more opportunities for smaller investors to analyze firm performance and to play a monitoring role in corporate governance. More individual investors can execute the power of shareholders at a lower cost by utilizing the tool of XBRL and monitor the wrong-doings of managers, adding "additional perceived pressures" on managerial decisions.

The adoption of XBRL not only favors smaller investors, but also facilitates

[7] Small investors include individual investors, small groups of investors, web-based or social connections-based investment clubs, and any other investment groups that have limited resources and similar risk preferences.

sophisticated users' ability to access, extract, and analyze firm performance data more efficiently and effectively. Sophisticated information users, such as institutional investors and analysts, are able to leverage their superior knowledge to obtain greater benefits from XBRL and enhance their information advantages (Blankespoor et al. 2014). In the pre-XBRL era, sophisticated information users bore the cost of information mining through self-supported agents (applications, software, and other programming-based intelligence macros). Searching, extracting, and formalizing data from complex, diversified financial reports takes resources away from analyzing the disclosed financial details. The XBRL mandate eliminates the need to convert financial information into machine-readable records, giving sophisticated users more time to conduct the value-added analyses. For instance, Bloomberg consumes XBRL data to fast-track company financials to analysts and has increased the usage every year including the footnotes which may not be captured in the past (Efthimides 2017). Liu et al. (2014) also document that XBRL adoption increases analyst coverage, and improves the timeliness and enhances the accuracy of analysts' forecasts. Consequently, by shifting more resources from tedious information collection to information analyses, larger investors may improve their ability to evaluate firm and management performance.

Financial reporting is an important source of firm-specific information for investors to monitor managers (Biddle et al. 2009). With expanded information accessibility and improved analytical ability, information from financial reports can be better accessed, extracted and interpreted to monitor managers' investment decisions. Enhanced external monitoring could mitigate the likelihood that managers invest in suboptimal projects due to the increased probability of being detected. For instance, Ferreira and Matos (2008) and

Chang et al. (2009) provide evidence that firms with improved monitoring (higher institutional ownership and greater analyst coverage) are less likely to make inefficient investments. By improving both smaller and sophisticated investors' ability to access, extract, and analyze financial reporting information, the adoption of XBRL reduces information asymmetry and strengthens external monitoring. Thus, the XBRL mandate may curb managers' suboptimal investment decisions.

Based on the issues discussed above, this essay proposes the following hypothesis:

H1: The adoption of XBRL has a positive impact on corporate investment efficiency.

The adoption of XBRL gives both small investors and sophisticated investors better access to firm-specific information, resulting in an improved monitoring mechanism and diminished information asymmetry. The potential enhanced monitoring mechanism may project "perceived pressure" on managers, prompting them to avoid inefficient investment behaviors. The effect of the XBRL mandate should be more evident for firms with weaker external monitoring, because managers in those firms have greater incentives to make suboptimal investments due to the lower odds of detection with ineffective monitoring. Thus, managers in these firms may be more inclined to adjust their misbehaviors after the XBRL-induced enhancement in monitoring. By contrast, if the existing governance or external monitoring mechanisms can ensure that investors gather sufficient information and conduct reasonable monitoring of management, the incremental gains from the adoption of XBRL will be reduced. Prior literature documents that the effectiveness of monitoring is positively related to the percentage of shares held by institutional investors and the number of analyst following (Kim et al. 2012; Khurana and Moser 2012).

Following prior studies, this study uses the percentage of institutional holdings and analyst coverage as proxies for the external monitoring function, and postulates the following hypothesis:

H2: The effect of XBRL adoption on investment efficiency is stronger for firms with weaker external monitoring.

In an uncertain information environment, the information for decision-making requires a high degree of aggregation and abstraction to produce manageable mappings. Prior literature suggests that a more complex environment increases the difficulty for investors to assess firms' opportunistic behaviors, such as earning management (Lim et al. 2008).

To be more specific, if a company is diversified with multiple business segments, firm-specific information is more difficult to track and aggregate in a timely manner. It is also more difficult for investors to infer underlying risks and values based on existing information. Thus, the problems created by information asymmetry will be more pronounced in a company with multiple business segments, resulting in a higher likelihood for managers to make inefficient investment decisions. Also, the existence of multiple business segments may diffuse monitoring resources, which lead to weaker monitoring of managerial opportunism. In addition, analyst forecast errors, which is the gap between the announced earnings and analysts' expectation, can be used to measure information uncertainty (Kim et al. 2012). In a more uncertain environment, it is more difficult for analysts to estimate reliable predictions as well. Prior literature also shows that information incorporated in analyst forecasts has significant information value to investors (e.g., Weiss 2010). Investors may not able to rely on information from analysts or other information intermediaries if they are inaccurate. Therefore, the benefits of the adoption XBRL on investment efficiency can be highlighted in an uncertain information environment.

Based on the discussions above, this essay proposes the following hypotheses:

H3: The effect of XBRL adoption on investment efficiency is stronger for firms with more uncertain information environment.

Apart from the complexity of companies themselves, this study also examines how the readability of financial reporting will affect the benefits of XBRL adoption on investment efficiency. Linguistic complexity increases the difficulty in consuming financial reports and leads to a wider information gap between investors and companies (Hoitash and Hoitash 2017). As we discussed before, successful communication between managers and investors is determined by the combination of the quality of the information and the processing cost. If the original financial reports have complicated semantics, words, and sentences, the interactive XBRL tagging systems will be less able to improve information processing efficiency due to the limited quality of financial information. By contrast, with more readable financial reports, the benefits of reduced processing cost will be magnified. Consistent with prior literature (Lehavy et al. 2011; Bozanic and Thevenot 2015), this research adopts two measures of linguistic complexity: the commonly used Fog Index (Gunning 1952) and the length of reports (Li 2008). Therefore, this study posits:

H4: The effect of XBRL adoption on investment efficiency is stronger for firms with higher readable financial reports.

4.3. Sample Selection and Empirical Specifications

4.3.1 Sample Selection

We collect the XBRL filings from EDGAR's database of Interactive Data Filings and Really Simple Syndication (RSS) feeds, and identify their initial (first) XBRL filing date as the adoption date. Then we focus only on firms that belong to Phase I, II or III filers[8] by limiting the sample to firms that adopted XBRL between June 16, 2009, and June 14, 2012, resulting in 22,353 firm-year observations merged with the Compustat database from the fiscal year 2008 to 2012. To avoid a potential self-selection bias in the sample, all XBRL filings issued by voluntary adopters were excluded from our sample.

Since our empirical design treats the adoption of XBRL as an exogenous shock, the initial date of mandatory XBRL filing is used to classify the pre and post period for each company. The final sample period runs from 2008 to 2012 to include both pre and post periods (Dong et al. 2016)[9]. Firm-specific fundamentals, stock price, and analyst forecast data were obtained from Compustat, the Center for Research in Security and Prices (CRSP), and the Institutional Brokers' Estimate Systems (I/B/E/S) respectively. After eliminating firms in the financial industry (SIC codes 6000-6999), and observations with missing data, our financial sample comprises 10,111 observations for 2,396 firms. Table 7 presents the summary of sample selection and Table 8 displays the descriptive statistics.

[8] The detail XBRL phase-in schedule can be found at: https://www.sec.gov/info/smallbus/secg/interactivedata-secg.htm

[9] The sample period is from 2008 to 2010 for phase I filers, from 2009 to 2011 for phase II filers, and from 2010 to 2012 for phase III filers.

Table 7 Sample Selection

	# of Obs.
Initial: Firm-years of XBRL adopters between 2009/6/15 to 2012/6/14	22,353
Less: Filings from voluntary XBRL adopters	(66)
Less: Filings from financial institutions (SIC 6000-6999)	(6,777)
Less: Filings that are not available for investment efficiency	(2,516)
Less: Filings that are missing control variables	(2,883)
Full Sample:	10,111

Table 7 provides details on the sample selection process.

Table 8 Descriptive Statistics

Variables	Full sample						XINV>=0			XINV<0		
	N	mean	median	p25	p75	sd	N	mean	median	N	mean	median
AbsXINV	10,111	0.187	0.109	0.048	0.218	0.282	3,840	0.229	0.101	6,271	0.161	0.113
Post	10,111	0.331	0	0	1	0.471	3,840	0.328	0	6,271	0.332	0
LnMve	10,111	5.977	6.170	4.313	7.582	2.296	3,840	5.998	6.201	6,271	5.965	6.151
Loss	10,111	0.341	0.000	0.000	1.000	0.474	3,840	0.312	0.000	6,271	0.359	0.000
Leverage	10,111	0.272	0.170	0.009	0.337	0.578	3,840	0.235	0.150	6,271	0.294	0.183
Cash	10,111	0.154	0.100	0.034	0.211	0.166	3,840	0.161	0.109	6,271	0.150	0.095
MTB	10,111	2.351	1.389	1.039	2.105	5.816	3,840	2.264	1.383	6,271	2.405	1.390
Analyst	10,111	7.760	5.000	0.000	12.000	9.042	3,840	7.795	5.000	6,271	7.739	4.000
StdCFO	10,111	0.102	0.050	0.028	0.092	0.232	3,840	0.102	0.053	6,271	0.101	0.048
StdSales	10,111	0.195	0.128	0.070	0.231	0.227	3,840	0.207	0.138	6,271	0.188	0.121
StdInvestment	10,111	0.276	0.111	0.050	0.259	0.488	3,840	0.275	0.132	6,271	0.276	0.099
Z-Score	10,111	0.153	-1.474	-2.628	-0.347	13.682	3,840	-0.438	-1.573	6,271	0.516	-1.425
Tangibility	10,111	0.264	0.177	0.075	0.386	0.243	3,840	0.246	0.153	6,271	0.275	0.191
OperCycle	10,111	2.766	3.835	1.773	4.600	2.633	3,840	2.710	3.822	6,271	2.800	3.851
FRQ	10,111	-0.047	-0.031	-0.054	-0.019	0.051	3,840	-0.046	-0.033	6,271	-0.047	-0.030

Table 8 presents descriptive statistics for the dependent, independent, and control variables included in our analysis. The sample runs from 2008 to 2012, resulting in 10,111 observations. The column XINV>=0 and XINV<0 indicates over- and under- investment groups. See Appendix B for variable definitions.

4.3.2 Empirical Specifications

Due to the unobservable nature of investment behaviors, investment efficiency cannot be measured directly by financial ratios. As suggested by prior literature (e.g., Biddle et al. 2009; Bae et al. 2016), efficient investment can be interpreted as a firm undertaking project with positive net present value (NPV). Under such interpretation, the under- or over-investment can be treated as either passing up investment opportunities with positive NPV or picking up projects with negative NPV. Following the same design philosophy as “abnormal accruals” or “abnormal audit fees” in the accounting literature (e.g., Kothari et al. 2005; Blankley et al. 2012), “abnormal investment” (XINV) can be measured as the magnitude (negative or positive) of deviations from the predicted optimal investment level based on firm-specific information. Specifically, positive “abnormal investment” (*XINV*>0) indicates an over-investment behavior, and negative “abnormal investment” (*XINV*<0) represents an under-investment behavior relative to the average investment level of peer firms in the same industry. The “abnormal investment efficiency” (*XINV*) is the residual value from an expected investment regression model, suggested by McNichols and Stubben (2008) as follows:

$$\begin{aligned} INV_{i,t} = \beta_0 + \beta_1 Q_{i,t-1} + \beta_2 Q_{i,t-1} \times Quartile2_{i,t-1} + \beta_3 Q_{i,t-1} \times Quartile3_{i,t-1} \\ + \beta_4 Q_{i,t-1} \times Quartile4_{i,t-1} + \beta_5 CF_{i,t} + \beta_6 Growth_{i,t-1} + \beta_7 INV_{i,t-1} \\ + \varepsilon_{i,t} \qquad (1) \end{aligned}$$

where *INV* is the investment level (capital expenditures), *Quartile2*, *Quartile3*, and *Quartile4* are indicator variables that equal to 1 if *Q* is in the second, third, and fourth quartiles of its industry-year distribution, *CF* is the cash flows, and *Growth* equals the natural log of total assets at the end of year t-1 divided by total assets at the end of year t-

2. The subscript "i" is the identifier for the firm, and the "t" is the representative of the year. The investment model is estimated separately for each industry and year.

To investigate the effect of XBRL adoption on corporate investment efficiency (H1), we estimate the following regression:

$$AbsXINV_{i,t} = \beta_0 + \beta_1 Post_{i,t} + CONTROLS_{i,t} + FIXED\ EFFECTS + \varepsilon_{i,t} \qquad (2)$$

where $AbsXINV_{i,t}$ is the absolute value of abnormal investment measured by equation (1) in year t. Our variable of interest is an indicator $Post_{i,t}$ that equals 1 if a firm-year is after the firm's XBRL adoption and 0 otherwise. To eliminate concerns arising from the gap between the adoption date and fiscal year end, we treat the variable $Post_{i,t}$ equals to 0 if the gap between the XBRL adoption date and the fiscal year end is less than 180 days. For instance, if a firm's fiscal year-end is on September 30, 2010, and the initial date of XBRL adoption is August 1, 2010, the gap between these two dates is less than 180 days, so $Post_{i,t}$ of this firm at the year 2010 equals to 0. This criterion ensures the number of XBRL adoption days is long enough to affect managerial incentives and consequently impact corporate policies[10].

CONTROLS represents a set of control variables adopted from prior investment efficiency studies (e.g., Bae et al. 2016; Biddle et al. 2009; Cheng et al. 2013; Lara et al. 2016). Specifically, we control firm size (*LnMve*), leverage (*Leverage*), cash (*Cash*), cash flow volatility (*StdCFO*), investment volatility (*StdInvestment*), financial distress risk (*Z-score*), tangible assets (Tangibility), operating cycle (*OperCycle*), and previous investment efficiency (*LagXINV*) as firm-specific characteristics. Since analyst coverage (*Analyst*) and

[10] In an untabulated test, we observe similar results if the 180-day adjustment is not applied.

sales volatility (*StdSales*) are identified to be related to firm investment efficiency, we also include them as controls. Additionally, we add firm performance such as operational earnings (*Loss*) and market to book value (*MTB*) as control variables. Since financial reporting quality improves investment efficiency (e.g., Biddle and Hilary 2006; Gomariz and Ballesta 2014), the financial reporting quality measurement (*FRQ*) is included as control variables as well[11]. Finally, we cluster standard errors by firm level to control within-firm correlations of residuals (Petersen, 2009). Appendix B provides detailed variables' definition.

To test a set of hypothesis that aims to investigate the moderating effect of external monitoring (H2), environmental uncertainty (H3), and financial reporting readability (H4), we employ the following regression:

$$AbsXINV_{i,t} = \beta_0 + \beta_1 Post_{i,t} + \beta_2 MODERATOR_{i,t} + \beta_3 Post_{i,t} \times MODERATOR_{i,t} + CONTROLS + FIXED\ EFFECTS + \varepsilon_{i,t} \quad (3)$$

where *MODERATOR* is one of the three sets of measures that captures the external monitoring, the environmental uncertainty, and the readability of reporting.

Following Lara et al. (2016) and Bae et al. (2016), we utilize institutional holding and analyst coverage to evaluate the monitoring functionality from external information users: institutional investors (*InstHold*) and analysts following (*Analyst*). We expect that the coefficients of the interaction terms are positive, suggesting that the effect of XBRL adoption is more pronounced for firms with weaker external monitoring.

---

[11] We adopt the definition of accruals quality from Francis et al. (2005), which is based on the Dechow and Dichev (2002) model. Additionally, we observe similar results when using discretionary accruals from Kasznik (1999) as an alternative measure.

The second set of moderators measure environmental uncertainty with the number of business segments (*NBus*) and analyst forecast errors (*Ferror*) (Dong et al. 2016; Kim et al. 2012). The variable of interest (interaction terms) captures the incremental effect of XBRL adoption when information uncertainty changes. We anticipate observing a negative association, implying an enhanced effect of XBRL adoption for firms with higher information uncertainty.

Finally, to test the fourth hypothesis, we substitute the moderator with the Fog Index (*FogIndex*) and the length of reporting (*Length*), which reflect the financial reporting readability (Li 2008; Hoitash and Hoitash 2017). As suggested by the hypothesis that more readable financial statements magnify the effect of the benefits of utilizing XBRL, we expect to observe a positive association between abnormal investment and the interaction terms.

4.4. Empirical Results

4.4.1 Main Results

Table 9 presents our main findings for the test of Hypothesis 1, examining the impact of the XBRL adoption on investment efficiency. The dependent variable is the absolute value of the abnormal investment (*AbsXINV*) derived from the expected investment model suggested by McNichols and Stubben (2008). The variable of interest is the dummy variable *Post*, indicating whether the firm has already stepped into the post-XBRL adoption period. The coefficient is -0.029 ($p<0.01$), suggesting that the adoption of XBRL will significantly reduce managers' suboptimal investments. Since the mean of abnormal investment is 0.187, the magnitude of reduction from XBRL mandate is around 16 percent. With respect to control variables, firms with higher operation volatilities in terms of cash flow, sales, and investment have inferior investment efficiency. Additionally, higher investment efficiency is associated with lower bankruptcy risk, higher analyst coverage, lower leverage, lower tangibility, and fewer prior capital investments. Moreover, we find consistent evidence that higher investment efficiency is associated with higher financial reporting quality. All coefficients of controls are consistent with prior studies.

Table 9 Main Results

Independent Variables	Dependent Variable
	AbsXINV
Intercept	0.249**
	(2.29)
Post	-0.029***
	(-3.22)
LnMve	-0.001
	(-0.51)
Loss	0.010
	(1.40)
Leverage	-0.001
	(-0.03)
Cash	0.085***
	(3.05)
MTB	0.004*
	(1.93)
Analyst	-0.000
	(-1.03)
StdCFO	0.076**
	(2.10)
StdSales	0.028
	(1.24)
StdInvestment	0.017*
	(1.85)
Z-Score	-0.001
	(-0.77)
Tangibility	-0.096***
	(-4.46)
OperCycle	-0.001
	(-0.49)
FRQ	-0.553***
	(-4.52)
LagXINV	0.176***
	(8.17)
Year and Industry fixed effects	Included
N	10,111
Adj-R^2	0.177

Table 9 reports the association between XBRL adoption and firms' investment efficiency (H1). The dependent variable (AbsXINV) represents the level of abnormal investments, and the residual derived from the equation following McNichols and Stubben (2008). Reported t-statistics are in parentheses and are estimated with the clustered standard error by firms as suggested by Petersen (2009). *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.4.2 External Monitoring

Table 10 reports the moderate effect of external monitoring on the relationship between the XBRL mandate and investment efficiency. In Panel A, the variable of interest is the interaction term of the number of shares institutional investors hold (InstHold) and the dummy variable *Post*. Institutional holdings and other control variables are included as well. We find a positive and significant coefficient (0.042, $p<0.05$), suggesting that the benefits of XBRL adoption on investment efficiency are mitigated by superior external monitoring. The coefficient of institutional holdings itself is positively significant at the 0.05 level, which is consistent with the monitoring role of institutional investors on suboptimal investments. In Panel B, the external monitoring is measured as the analyst following (Analyst). Consistent with the results in panel A, the coefficient of the interaction term is positive and significant (0.001; $p<0.01$). Overall, the results are in line with our second hypothesis that the benefits of adopting the XBRL in improving investment efficiency will be mitigated by the existence of stronger external monitoring.

Table 10 The Impact of External Monitoring on the Association between the XBRL Mandate and Investment Efficiency

Panel A Institutional Holdings

Independent Variables	Dependent Variable
	AbsXINV
Intercept	0.255**
	(2.36)
*Post*InstHold*	0.042**
	(2.44)
InstHold	-0.043***
	(-4.19)
Post	-0.057***
	(-4.09)
LnMve	0.001
	(0.64)
Loss	0.009
	(1.30)
Leverage	-0.002
	(-0.09)
Cash	0.080***
	(2.81)
MTB	0.004*
	(1.84)
Analyst	-0.000
	(-0.78)
StdCFO	0.074**
	(2.04)
StdSales	0.026
	(1.15)
StdInvestment	0.015
	(1.57)
Z-Score	-0.001
	(-0.65)
Tangibility	-0.099***
	(-4.49)
OperCycle	-0.001
	(-0.31)
FRQ	-0.540***
	(-4.37)
LagXINV	0.180***
	(8.05)
Year and Industry fixed effects	Included
N	9,668
Adj-R^2	0.180

Panel B Analyst Coverage

Independent Variables	Dependent Variable
	AbsXINV
Intercept	0.253**
	(2.33)
*Post*Analyst*	0.001***
	(2.72)
Analyst	-0.001***
	(-2.69)
Post	-0.043***
	(-4.12)
LnMve	-0.001
	(-0.33)
Loss	0.010
	(1.38)
Leverage	-0.001
	(-0.04)
Cash	0.086***
	(3.06)
MTB	0.004*
	(1.92)
StdCFO	0.076**
	(2.10)
StdSales	0.028
	(1.24)
StdInvestment	0.017*
	(1.87)
Z-Score	-0.001
	(-0.76)
Tangibility	-0.096***
	(-4.45)
OperCycle	-0.001
	(-0.50)
FRQ	-0.548***
	(-4.49)
LagXINV	0.175***
	(8.17)
Year and Industry fixed effects	Included
N	10,111
Adj-R^2	0.177

Table 10 presents the result for H2 by interacting the proxy for external monitoring (institutional holdings in Panel A and analyst coverage in Panel B) with the XBRL adoption dummy Post. The dependent variable (AbsXINV) represents the level of abnormal investments,and the residual derived from the equation following McNichols and Stubben (2008). Reported t-statistics are in parentheses and are estimated with the clustered standard error by firms as suggested by Petersen (2009). *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.4.3 Environment Uncertainty

Table 11 shows the results of the incremental effect of environmental uncertainty on the association between XBRL adoption and investment efficiency (H3). In Panel A, we examine how the firm's operational complexity (NBus) will affect the benefits of XBRL in improving investment efficiency. Our expectation is that such benefits will be magnified by the firm's operational complexity, since XBRL can better serve investors to track, extract, and aggregate information for complex firms with multiple business segments. Indeed, the coefficient is significantly negative (-0.009, p<0.01), supporting the hypothesis. Panel B displays the findings for the alternative dimension of environmental uncertainty, analyst forecast errors (Ferror). A higher analyst forecast error indicates that a firm's operating performance is full of uncertainty and hard to predict. In such an inferior information environment, investors can hardly generate reliable and consensus opinions on the evaluation of the firm's investment activities. Consistent with our expectation, the coefficient of the interaction term between analyst forecast errors and *Post* is significantly negative (-0.153, p<0.05). Collectively, our findings provide evidence that the improvement in investment efficiency after the XBRL adoption is greater for firms with higher environmental uncertainty.

Table 11 The Impact of Environmental Uncertainty on the Association between the XBRL Mandate and Investment Efficiency

Panel A Number of Business Segments

Independent Variables	Dependent Variable *AbsXINV*
Intercept	0.204* (1.87)
*Post*NBus*	-0.009*** (-2.79)
NBus	0.001 (0.29)
Post	-0.010 (-0.84)
LnMve	-0.000 (-0.23)
Loss	0.010 (1.33)
Leverage	-0.001 (-0.04)
Cash	0.082*** (2.81)
MTB	0.004* (1.88)
Analyst	-0.000 (-0.94)
StdCFO	0.074** (2.03)
StdSales	0.026 (1.07)
StdInvestment	0.017* (1.76)
Z-Score	-0.001 (-0.72)
Tangibility	-0.099*** (-4.32)
OperCycle	-0.001 (-0.43)
FRQ	-0.552*** (-4.36)
LagXINV	0.179*** (8.08)
Year and Industry fixed effects	Included
N	9,463
Adj-R^2	0.178

Panel B Analyst Forecast Errors

Independent Variables	Dependent Variable
	AbsXINV
Intercept	0.266*
	(1.85)
*Post*Ferror*	-0.153**
	(-2.52)
Ferror	0.032
	(0.72)
Post	-0.012
	(-1.16)
LnMve	-0.003
	(-1.40)
Loss	0.009
	(1.15)
Leverage	0.043*
	(1.76)
Cash	0.134***
	(3.27)
MTB	0.009***
	(2.59)
Analyst	0.000
	(0.08)
StdCFO	0.121*
	(1.66)
StdSales	0.020
	(0.65)
StdInvestment	0.013
	(1.10)
Z-Score	-0.005***
	(-2.86)
Tangibility	-0.066***
	(-3.89)
OperCycle	0.000
	(0.15)
FRQ	-0.272*
	(-1.93)
LagXINV	0.146***
	(6.45)
Year and Industry fixed effects	Included
N	6,705
Adj-R^2	0.146

Table 11 provides results for testing H3 by interacting environmental uncertainty proxies (operational complexity in Panel A and analyst forecast error in Panel B) with XBRL adoption. The dependent variable (AbsXINV) represents the level of abnormal investments, and the residual derived from the equation following McNichols and Stubben (2008). Reported t-statistics are in parentheses and are estimated with the clustered standard error by firms as suggested by Petersen (2009). *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.4.4 Financial Reports Readability

Table 12 presents the influence of reports readability on the association between XBRL adoption and investment efficiency. In Panel A, we use the commonly used readability measure, the Fog Index (FogIndex), to proxy for the readability of financial reporting. The results indicate that the coefficient of the interaction term between the Fog Index and the *Post* is significantly positive (0.008, $p<0.05$), implying that the benefits of adopting XBRL are mitigated when the original financial filings are less readable. Additionally, in Panel B, we replace the Fog Index with the length of reporting (Length) and find a consistent result (0.041, $p<0.01$) of the moderate effect of readability on the association between XBRL adoption and investment efficiency. The above findings are consistent with our expectation (H4) that investors will receive increased benefits from the use of XBRL to extract and interpret information when the financial report is more readable.

Table 12 The Impact of Readability on the Association between the XBRL Mandate and Investment Efficiency

Panel A Fog Index

Independent Variables	Dependent Variable
	AbsXINV
Intercept	0.293*** (2.61)
*Post*FogIndex*	0.008** (2.12)
FogIndex	-0.002 (-1.35)
Post	-0.211** (-2.44)
LnMve	-0.001 (-0.50)
Loss	0.010 (1.36)
Leverage	-0.003 (-0.15)
Cash	0.093*** (3.16)
MTB	0.004* (1.84)
Analyst	-0.000 (-0.86)
StdCFO	0.075** (2.02)
StdSales	0.030 (1.29)
StdInvestment	0.014 (1.42)
Z-Score	-0.001 (-0.63)
Tangibility	-0.091*** (-4.05)
OperCycle	-0.001 (-0.34)
FRQ	-0.540*** (-4.22)
LagXINV	0.186*** (8.22)
Year and Industry fixed effects	Included
N	9,457
Adj-R^2	0.180

Panel B The length of reporting

Independent Variables	Dependent Variable *AbsXINV*
Intercept	0.485***
	(3.48)
*Post*Length*	0.041***
	(3.23)
Length	-0.023***
	(-2.63)
Post	-0.470***
	(-3.45)
LnMve	0.001
	(0.22)
Loss	0.012
	(1.63)
Leverage	-0.002
	(-0.10)
Cash	0.091***
	(3.13)
MTB	0.004*
	(1.75)
Analyst	-0.000
	(-1.15)
StdCFO	0.074**
	(2.00)
StdSales	0.031
	(1.34)
StdInvestment	0.015
	(1.51)
Z-Score	-0.001
	(-0.60)
Tangibility	-0.092***
	(-4.13)
OperCycle	-0.001
	(-0.40)
FRQ	-0.534***
	(-4.18)
LagXINV	0.185***
	(8.22)
Year and Industry fixed effects	Included
N	9,457
Adj-R^2	0.181

Table 12 provides empirical evidence is supportive of H4 by interacting financial reporting complexity (Gunning's Fog Index in Panel A and the length of financial reporting in Panel B) with XBRL adoption. The dependent variable (AbsXINV) represents the level of abnormal investments, and the residual derived from the equation following McNichols and Stubben (2008). Reported t-statistics are in parentheses and are estimated with the clustered standard error by firms as suggested by Petersen (2009). *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.5. Additional Tests

4.5.1 Non-capital Investments

In prior sections, we conduct analyses on a firm's capital expenditures. As a robustness check, we use non-capital investments to examine the changes in investment efficiency. We compute non-capital expenditures as the sum of R&D expenditures and acquisitions, scaled by lagged total assets. We use the same expected investment model to obtain abnormal investment levels and re-run our main model. The results are reported in Table 13.

The main effect of XBRL adoption (*Post*) is still negative and significant (-0.432, $p<0.01$), suggesting that the XBRL mandate improves non-capital investment efficiency as well. Thus, the above results are consistent with our main findings and show the lower information processing cost from the XBRL mandate enhances investment efficiency for both capital and non-capital investments.

Table 13 Additional Analysis Investigating the XBRL Benefits on Non-Capital Investments

Independent Variables	Dependent Variable
	AbsNXINV
Intercept	0.328
	(1.26)
Post	-0.423***
	(-3.69)
LnMve	0.055***
	(2.64)
Loss	0.137
	(1.62)
Leverage	0.139
	(0.56)
Cash	-0.448
	(-1.18)
MTB	0.032
	(1.59)
Analyst	-0.002
	(-0.51)
StdCFO	0.757**
	(2.39)
StdSales	-0.221
	(-0.93)
StdLInvestment	0.091***
	(4.89)
Z-Score	-0.009
	(-0.55)
Tangibility	-1.709***
	(-9.13)
OperCycle	-0.065**
	(-2.44)
FRQ	-3.621***
	(-2.67)
LagXLINV	0.281***
	(9.26)
Year and Industry fixed effects	Included
N	10,106
Adj-R^2	0.367

Table 13 presents results for additional tests on the relationship between XBRL adoption and firms' non-capital investment efficiency. The dependent variable (AbsXINV) represents the level of abnormal investments, and the residual derived from the equation following McNichols and Stubben (2008). Reported t-statistics are in parentheses and are estimated with the clustered standard error by firms as suggested by Petersen (2009). *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.5.2 Difference-in-Difference Design

In this section, a difference-in-difference design is applied as a robustness test to compare the effect of XBRL on adopters versus non-adopters in pre-adoption versus post-adoption periods. We build two subsamples based on the phase of XBRL adoption. In first (second) subsample, the treatment groups contain phase I (II) filers, and phase II (III) filers are included as the control group. The sample period is one year before and one year after the XBRL adoption of treatment firms in each subsample. Following prior studies, we estimate the following OLS regression model:

$$AbsXINV_{i,t} = \beta_0 + \beta_1 After_{i,t} + \beta_2 XBRL_{i,t} + \beta_3 After * XBRL_{i,t} + CONTROLS_{i,t} + FIXED\ EFFECTS + \varepsilon_{i,t} \quad (4)$$

where *AFTER* is a time indicator that equals to 1 if firms are at treatment firms' post-adoption period, and 0 otherwise. *XBRL* is a firm indicator that equals to 1 if firms belong to the treatment group, and 0 otherwise.

Column 2 of Table 14 present the results of DID models for the first subsample (phase I filers vs. phase II filers). However, the coefficient on *AFTER*XBRL* is not significant, suggesting the effect of XBRL adoption is weak for phase I filers. A possible explanation is that phase I filers are large firms and already have relatively strong governance and efficient investment. The regression results of subsample two (phase II adopter vs. phase III adopters) are provided in column 3 of table 14. The coefficient on *After*XBRL* is negative and significant (-0.038, p<0.1), further confirming our main findings.

Table 14 Difference-in-Difference Design

Independent Variables	Phase1 vs. Phase2	Phase2 vs. Phase3
	Dependent Variable	Dependent Variable
	AbsXINV	*AbsXINV*
Intercept	0.134	-0.108
	(1.40)	(-1.19)
*After*XBRL*	0.037	-0.038*
	(1.17)	(-1.64)
XBRL	0.001	-0.030
	(0.07)	(-1.42)
After	-0.001	0.044**
	(-0.04)	(2.08)
LnMve	-0.009	0.012
	(-0.98)	(0.89)
Loss	-0.006	0.004
	(-0.43)	(0.19)
Leverage	0.026	0.091
	(0.33)	(1.41)
Cash	-0.100	0.103
	(-0.71)	(0.97)
MTB	0.033	0.013*
	(1.29)	(1.81)
Analyst	0.001	-0.000
	(0.90)	(-0.48)
StdCFO	-0.138	0.057
	(-0.47)	(0.44)
StdSales	-0.015	-0.039
	(-0.35)	(-1.00)
StdInvestment	0.088	0.014
	(1.40)	(0.62)
Z-Score	0.008	-0.009*
	(0.51)	(-1.78)
Tangibility	-0.028	-0.051
	(-0.63)	(-1.21)
OperCycle	-0.003	0.005
	(-0.63)	(1.05)
FRQ	-0.097	-0.560
	(-0.38)	(-1.21)
LagXINV	0.002	0.246***
	(0.02)	(4.36)
Year and Industry fixed effects	Included	Included
N	466	1263
Adj-R^2	0.156	0.176

Table 14 presents results for additional tests on the relationship between XBRL adoption and firms' capital investment efficiency based on a difference-in-difference design. Columns 1 provides results based on Phase I and Phase II filers from 2009 to 2010. Columns 2 shows results based on Phase II and Phase III filers from 2010 to 2011. The dependent variable (AbsXINV) is measured following McNichols and Stubben (2008). Reported t-statistics are based on robust standard errors adjusted using a cluster at the firm-year level. *, **, *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.5.3 Change Design

Although we have included the control variables in the main model, we cannot rule out unobservable factors affecting managers' decision on investment. Bias caused by unobservable characteristics can be eliminated using a change model specification if the unobservable differences remain invariant during the period of study (Lennox et al. 2011). Therefore, a year-to-year change research design is employed to further confirm the relation between investment efficiency and XBRL adoption.

$$D.AbsXINV_{i,t} = \beta_0 + \beta_1 D.Post_{i,t} + D.CONTROLS_{i,t} + FIXED\ EFFECTS + \varepsilon_{i,t} \quad (5)$$

where all variables in the model (5) are the first-order difference of variables from the model (2).

Table 15 shows the results of this model. Consistent with our expectation, the coefficient of *D.Post* is negative and significant (-0.017, p<0.05), which further confirms the effect of XBRL adoption on investment efficiency after considering the unobservable effects.

Table 15 Change Design

Independent Variables	Dependent Variable
	D.AbsXINV
Intercept	0.306
	(1.60)
D.Post	-0.165**
	(-2.03)
D.LnMve	0.037***
	(3.84)
D.Loss	-0.007
	(-0.81)
D.Leverage	-0.097***
	(-3.18)
D.Cash	-0.105*
	(-1.86)
D.MTB	-0.002
	(-0.46)
D.Analyst	0.004***
	(3.41)
D.StdCFO	0.138**
	(2.18)
D.StdSales	0.043
	(0.77)
D.StdInvestment	-0.239***
	(-6.16)
D.Z-Score	0.004**
	(2.04)
D.Tangibility	0.638***
	(4.71)
D.OperCycle	0.007
	(0.75)
D.FRQ	-0.456**
	(-2.22)
D.LagXINV	-0.347***
	(-16.95)
Year and Industry fixed effects	Included
N	7,637
Adj-R^2	0.289

Table 15 presents results for additional tests on the relationship between XBRL adoption and firms' capital investment efficiency based on a change design. The dependent variable (D.AbsXINV) is the first order difference in abnormal investment (AbsXINV), which is measured following McNichols and Stubben (2008). Reported t-statistics are based on robust standard errors adjusted using a cluster at the firm-year level. *, **, *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.5.4 Dynamic Effect

In this section, we investigate the dynamic effect of XBRL adoption on investment efficiency. The successful adoption of XBRL is built on technological advancement and the familiarity of utilizing such technology by information users. We believe there will be a learning effect for information users to accumulate experience from maximizing the benefits of XBRL on lowering information processing cost. Consequently, consistent with the prior literature (Dong et al. 2016), we expect the impact of XBRL adoption on investment efficiency enhances as time goes. We modify equation (2) and estimate the following regression:

$$AbsXINV_{i,t} = \beta_0 + \beta_1 Post1_{i,t} + \beta_2 Post2_{i,t} + \beta_3 Post3_{i,t} + CONTROLS_{i,t} + FIXED\ EFFECTS + \varepsilon_{i,t} \quad (6)$$

where the *Post* is replaced with three event-year indicators, *Post1*, *Post2*, and *Post3*. *Post1* (*Post2* or *Post3*) equals 1 for observations that are 1 year (2 or 3 years) after a firm's XBRL adoption and 0 otherwise.

In table 16, we find that the coefficients of the three post-adoption dummies are negative and significant (-0.015, p<0.1 for *Post1*; -0.027, p<0.05 for *Post2*; -0.040, p<0.01 for *Post3*), demonstrating that the effect of XBRL adoption on investment efficiency is persistent after the adoption. More importantly, we find an increasing effect of the XBRL adoption in term of both significance and magnitude, supporting our expectation that investors face a learning curve in understanding XBRL and the effect of XBRL adoption on investment efficiency enhances as time goes.

Table 16 Dynamic Effects

Independent Variables	Dependent Variable
	AbsXINV
Intercept	0.174***
	(4.85)
Post1	-0.015*
	(-1.72)
Post2	-0.027**
	(-2.22)
Post3	-0.040***
	(-2.64)
LnMve	-0.004***
	(-2.62)
Loss	0.007
	(1.21)
Leverage	-0.020
	(-1.45)
Cash	0.128***
	(5.40)
MTB	0.002
	(1.55)
Analyst	0.000
	(0.51)
StdCFO	0.001***
	(3.23)
StdSales	0.059***
	(3.05)
StdInvestment	0.027***
	(3.00)
Z-Score	-0.001
	(-0.39)
Tangibility	-0.095***
	(-5.68)
OperCycle	-0.001
	(-0.48)
FRQ	-0.576***
	(-5.56)
LagXINV	0.168***
	(10.51)
Year and Industry fixed effects	Included
N	15,522
Adj-R^2	0.190

Table 16 provides results for the dynamic effects of XBRL adoption on firms' capital investment efficiency. The dependent variable (AbsXINV) represents the level of abnormal investments, and the residual derived from the equation following McNichols and Stubben (2008). Post1 (Post2 or Post3) equals 1 for observations that are 1 year (2 or 3 years) after a firm's XBRL adoption, and 0 otherwise. Reported t-statistics are in parentheses and are estimated with the clustered standard error by firms as suggested by Petersen (2009). *,

, and * indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.5.5 Over-Investment and Under-Investment

This additional test separates the effect of XBRL adoption on over-investment and under-investment. In Table 17 column 1, we only include observations of over-investment (*XINV*>0) and find negative and significant coefficient (-0.047, p<0.05) for *Post*, indicating that the lower information processing cost from the adoption of XBRL is associated with lower over-investment. In column 2, only observations of under-investment (*XINV*<0) are included in our model. Although negative, the coefficient of the Post dummy variable is insignificant, suggesting that the XRBL mandate is more likely to curb management opportunistic over-investment activities rather than under-investment.

The above results could be explained as follows. Since investors have no information about managers' potential project options, the evaluation of investment behaviors is based on the ex-post financial outcomes. Before managers invest in certain projects and make that investment information public, investors have no access to monitor their investment behaviors. Thus, it is much harder for investors to detect that managers took no actions and missed investment opportunities in the past. However, over-investment can be detected by investors easier since the investment costs and profits are recorded in companies' reports. Because the adoption of XBRL reduce the cost of processing financial information and facilitates the efficiency of detecting sub-optimal behaviors based on financial reports, managers may avoid more over-investments than under-investments. Additionally, investors' punishment for over-investment is much more severe than for under-investment. Under-investment may waste profitable opportunities to add firm values, but it does not erode shareholders' interests to benefit managers. Thus, shareholders are likely to punish managers' over-investment more than under-investment through market responses or monitoring mechanisms, and the managers are more likely to reduce over-investment than under-investment to avoid punishment from the market.

Table 17 Over-Investment and Under-Investment

Independent Variables	Dependent Variable: *AbsXINV*	
	XINV>=0	XINV<0
Intercept	0.094*	0.354**
	(1.73)	(2.25)
Post	-0.047**	-0.007
	(-2.41)	(-0.96)
LnMve	-0.006	0.000
	(-1.25)	(0.26)
Loss	0.007	0.014***
	(0.41)	(2.99)
Leverage	0.077	-0.022
	(1.28)	(-1.49)
Cash	0.219***	-0.010
	(3.28)	(-0.46)
MTB	0.002	0.005**
	(0.47)	(2.38)
Analyst	-0.001	0.000
	(-1.64)	(0.16)
StdCFO	0.061	0.071**
	(0.83)	(2.08)
StdSales	0.079*	-0.029
	(1.65)	(-1.47)
StdInvestment	0.033	0.013
	(1.41)	(1.63)
Z-Score	-0.004	0.000
	(-1.27)	(0.20)
Tangibility	-0.195***	-0.030
	(-5.13)	(-1.15)
OperCycle	-0.002	0.000
	(-0.53)	(0.13)
FRQ	-0.699***	-0.490***
	(-2.71)	(-3.77)
LagXINV	0.167***	0.183***
	(4.13)	(6.85)
Year and Industry fixed effects	Included	Included
N	3,840	6,271
Adj-R^2	0.168	0.313

Table 17 reports the association between XBRL adoption and firms' over-investment and under-investment. The dependent variable (AbsXINV) represents the level of abnormal investments, and the residual derived from the equation following McNichols and Stubben (2008). Reported t-statistics are in parentheses and are estimated with the clustered standard error by firms as suggested by Petersen (2009). *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, using two-tailed tests.

4.6 Conclusion

Market friction, inducing information asymmetry and agency conflicts, leads to suboptimal investment decisions. Extant prior literature documents that such imperfections can be constrained by financial disclosure. Then, a growing number of studies provide empirical evidence to support the link between the quality of financial reporting and investment efficiency. However, these studies mainly focus on the quality of the content in financial reports but overlook the importance of the cost of processing information. We fill the research gap by utilizing the adoption of XBRL as an exogenous shock to examine the effect of information processing efficiency on firms' investment behaviors.

Our results suggest that the adoption of XBRL can significantly reduce abnormal investments, especially over-investments. We further investigate potential factors that may magnify or mitigate the benefits of XBRL adoption on investment efficiency. Our results show that such benefits will be magnified in firms with inferior external monitoring, severer environment uncertainty, and more readable financial reports. These results are robust to a difference-in-difference (DID) and change research design, and continue to hold for non-capital investments. Considering the learning ability of investors in understanding XBRL, we also find a persistent and increasing pattern of the XBRL effect, supporting our expectation that the effect of XBRL adoption on investment efficiency enhances as time goes. Overall, our results contribute to the existing accounting literature by directly examining the effects of the efficiency of information processing on investment efficiency. The empirical setting is free of endogeneity and provides a clear interpretation by utilizing XBRL adoption as an exogenous shock.

This study is not without limitation. There is considerable overlap between the period of financial crisis and our sample period. Therefore, it is possible that our results are driven by the special financial conditions during the financial crisis. However, the results of the DID tests show that the effect of XBRL adoption on investment efficiency is stronger in the sample period away from the financial crisis, partially mitigating the concern.

CHAPTER 5: APPLYING ROBOTIC PROCESS AUTOMATION (RPA) IN AUDITING: A FRAMEWORK

5.1 Introduction

Traditional audit procedures are labor-intensive and time-consuming (Chan and Vasarhelyi 2011). To free human auditors from doing repetitive and low-judgment audit tasks and help them to focus on procedures requiring professional judgment, prior literature has proposed for decades that labor-intensive audit tasks be replaced with automation (e.g., Vasarhelyi 1984; Vasarhelyi and Halper 1991). One early application of automation technology in auditing is continuous auditing (CA). In recent years, commercial audit analytics software and electronic spreadsheets such as Microsoft Excel have been widely employed to automate tests and analyses. One advanced audit management system, an electronic workpaper system, also allows auditors to address specific client risks and communicate more efficiently with audit team members. Although technology has had a significant impact on improving audit efficiency, integration across multiple systems or applications is performed mainly by auditors, meaning that the actual external audit is still labor-intensive (Srinivasan 2016).

For this reason, practitioners have been interested in rethinking their processes in terms of automation and taking advantage of advanced automation technologies such as robotic process automation (RPA). RPA is a methodology that performs routine business processes by automating the way people interact, with multiple applications or analyses through a user interface and also by following simple rules to make decisions (Deloitte 2017). In the accounting domain, major accounting firms are considering applying RPA to achieve cost savings and increase operational efficiency. For instance, KPMG has recently announced that it will work with a global leader in enterprise RPA to help clients automate manual business processes (KPMG 2017). As one of the largest RPA consultants, Ernst and Young (EY) has delivered RPA projects to financial services organizations across 20 countries (EY 2016a). Although repetitive, structured and labor-intensive audit tasks (such as reconciliations, internal control testing, and detail testing) are ideal candidates for RPA, academic research into the application of RPA to auditing remains unexplored (Moffitt et al. 2018). If such audit tasks can be automated, auditors will be able to focus on tasks requiring professional judgments. Additionally, automated audit procedures will no longer be limited by the constrained processing power of human beings; consequently, the scale of the audit can be increased and more comprehensive audit evidence can be collected, enhancing audit quality. The purpose of this study is to propose a framework for applying RPA in audit practice and report the results of a pilot project to automate the confirmation process.

5.2. Background and Literature Review

5.2.1 Audit Procedure Automation

Because of its labor intensiveness and the range of decision structures, auditing has adopted automation technology for more than three decades (Issa et al. 2016). Vasarhelyi and Halper (1991) proposed the concept of continuous auditing (CA), which is defined as "*a methodology for issuing audit reports simultaneously with, or a short period of time after, the occurrence of the relevant events*" (CICA/AICPA 1999). Later, continuous auditing (CA) and continuous monitoring (CM) (hereafter CA/CM) became one of the applications of automation technology (Vasarhelyi et al. 2004). As part of their pilot project, Alles et al. (2008) apply CA/CM to an internal IT audit process. They develop guidelines for the formalization of the audit procedures into a computer-executable format and determine which procedures are automatable and which require reengineering. Many internal/IT audit procedures have been demonstrated to be automatable, thus saving costs, allowing for more frequent audits, and freeing up the audit staff for tasks that require human judgment (AICPA 2015).

Audit software vendors such as ACL, CaseWare, and CA technology, provide commercial and standardized IT packages and analytical software, supporting automation of basic audit tests such as three-way matching, sampling, and handling fairly large data sets (Appelbaum et al. 2017). Furthermore, commercial spreadsheets such as Microsoft Excel allow auditors to perform tests effectively and efficiently. Instead of manually filtering or copying and pasting spreadsheet data, the macro programming language Visual Basic Application (VBA) is widely employed to automate various tasks or analyses.

To further improve audit effectiveness and efficiency, accounting firms have adopted audit management systems such as electronic systems for workpaper preparation and review. An electronic workpaper system can enhance the audit quality by tailoring the file to address specific client risks, including setting the strategy to be used during the engagement and altering the nature, timing, and/or extent of planned audit procedures. In addition, electronic systems allow auditors to direct reference of information between documents and enables managers/reviewers to electronically access files and communicate remotely with their audit teams (Agoglia et al. 2010; Bedard et al. 2006).

Recent literature has suggested that auditors take advantages of emerging technologies to automate audit procedures. Yan et al. (2017) propose the Automated Contract Analysis System (ACAS) framework, which is based on auditing standards with contract-specific requirements. They demonstrate the feasibility, through the proposed ACAS framework, of incorporating text mining into contract audit procedures to automate contract analysis in the audit stages of risk assessment, substantive tests, and review, and to provide auditors with contract data that can be used to identify audit risk and generate audit evidence. Furthermore, Appelbaum and Nehmer (2017) propose the use of drones in audit automation and continuous auditing environments and illustrate how drones fit into audits for inventory count by gathering evidence to support specific assertions made by management.

5.2.2 Robotic Process Automation

By automating audit tests and analyses, technologies such as CA/CM, analytical tools, and electronic workpaper systems have significantly improved the audit effectiveness and efficiency. However, these technologies focus mainly on automating a specific task or test, leaving the coordination and integration across different systems or applications to be largely performed by auditors, reducing the improved audit efficiency and causing the actual audit to remain labor intensive. The application of an emerging technology in business automation, robotic process automation (RPA), which is an overlay for existing IT systems, may be a solution to the problem by executing a combination of audit tasks or analyses in multiple unrelated software systems.

The idea behind RPA is not new; it is traditional automation in terms of assembly line technology (Moffitt et al. 2018). RPA is defined as "*a preconfigured software instance that uses business rules and predefined activity choreography to complete the autonomous execution of a combination of processes, activities, transactions, and tasks in one or more unrelated software systems to deliver a result or service with human exception* *management*" (IEEE Corporate Advisory Group 2017, p11). An example of the RPA process is the retrieval of information from one system and entering the same information into another system or activating another system function. Unlike some traditional IT implementation and business reengineering that changes the existing systems, RPA tries not to disturb underlying IT systems and only replaces the existing manual process with the automated process through a presentation layer (IRPA 2016). Therefore, compared with major IT platform updates, the burdens of RPA implementation (costs, timelines, and risks) are relatively insignificant (EY 2016b).

RPA tools help businesses improve the efficiency of processes and the effectiveness of services. First, replacing the human workforce reduces the cost and processing time for high-frequency tasks. The running cost of an RPA software is around one-ninth that of employing a human being, and RPA "robots" can work 24 hours a day and 7 days a week (Burgess 2016). Secondly, the accuracy of the business process is improved. As long as an RPA tool is properly programmed, there is no need to worry that software robots will make the mistakes that human beings might (IRPA 2015). Finally, RPA offers flexibility and scalability. Once a process has been executed by a software robot, it can be scheduled for a particular time. In addition, the RPA robot is capable of performing many types of processes and can be quickly reassigned to other processes (Deloitte 2017). Figure 1 presents how RPA replicates a manual analysis-related task. Almost all steps, such as receiving email, activating functions in an ERP system, entering information into a spreadsheet, running analyses, and reporting results via email, are able to be performed by RPA, while only some exceptions are handled manually.

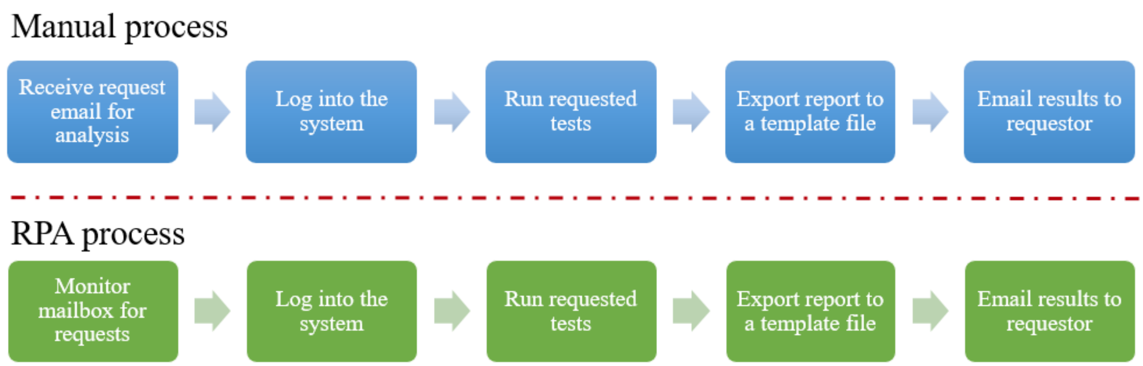


Figure 1 Manual Process vs. Robotic Process

Because of its low implementation cost and high potential benefits, RPA has been widely adopted in many industries. As of April 2015, Telefónica O2, the second-largest mobile telecommunications provider in the United Kingdom, had adopted more than 100 RPA "robots" to handle 500,000 transactions each month (Lacity et al. 2015). In addition, a business process outsourcing provider automated 14 core processes with RPA, achieving a typical 30% cost saving per process and improving service quality and accuracy. Also, in the process of updating London Premium Advice Notes (LPANs) to a central insurance market repository, an RPA robot was used to automate the most onerous steps: validating data, accessing the database, creating documents, and uploading the repository. After adoption, the processing time was only 30 minutes instead of several days (Deloitte 2017).

Even though the industry has observed the benefits of RPA, its applications in auditing practice are still unexplored. Additionally, many audit tasks are well-defined, highly repetitive, and predictable; for example, extracting exogenous information (confirmations from the electronic platform and customer reviews from social media) and matching information from multiple systems, which are multi-step tasks across multiple systems, are the ideal candidates for RPA (IRPA 2015). With the improved processing power of RPA, the scale of audit procedures can be increased and auditors will be able to focus on tasks that require professional judgment and higher order thinking skills, thereby enhancing audit quality. In this study, we propose a framework to apply RPA in auditing and demonstrate its feasibility by implementing a pilot project for the confirmation process.

5.3. Methodology

The framework proposed in this study was developed by following the design science approach (Gregor and Hevner 2013; Hevner et al. 2004). Design science is a research methodology that seeks the creation of new knowledge or understanding to solve real-world problems through the design of novel or innovative artifacts (e.g., things, processes, algorithms, and frameworks) and the evaluation of such artifacts (Hevner et al. 2004; Simon 1996).

This study develops a feasible framework for using an advanced automation technique (i.e., RPA) to address auditors' needs. Based on a review of the literature and professional guidelines for the framework for data analytics as well as prior studies addressing the use of RPA in the audit process (e.g., AICPA 2017; Moffitt et al. 2018; No et al. 2018), we identified four steps for effectively applying RPA to audit procedures. Finally, the framework was then evaluated by discussions with researchers and auditors and a pilot RPA implementation.

5.4. The Description of the Framework

Figure 2 illustrates the four steps, or stages, in the framework: 1) procedure selection, 2) data understanding, 3) RPA implementation, and 4) feedback and evaluation. Following is a more detailed description of each stage.

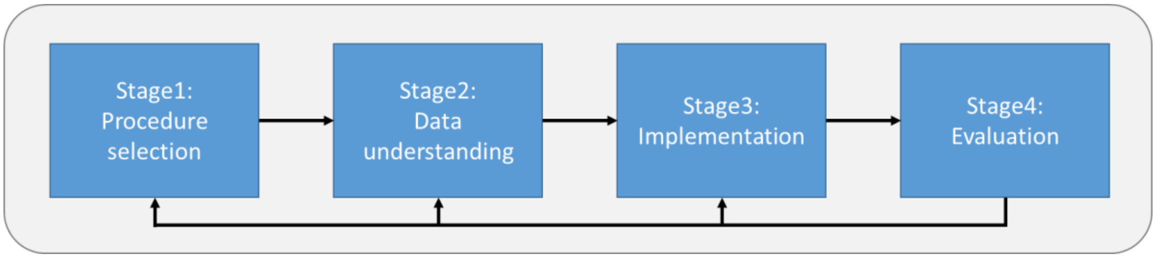


Figure 2 The Framework

Stage1: Procedure selection

When planning the application of RPA, accounting firms should review the structure of their audit procedures and identify appropriate candidates based on several criteria. First, well-defined audit procedures are more appropriate for RPA, because RPA software needs explicit instructions to complete tasks. Second, the tasks should be highly repetitive. Third, mature audit tasks should be automated first because the outcomes and cost are more predictable (Lacity et al. 2015). Abdolmohammadi (1999) considers audit tasks as structured if they are well defined and require very little judgment, while tasks with many alternative solutions that require considerable judgment are regarded as unstructured. Semi-structured tasks, which have limited alternative solutions and require a medium level of judgment, fall somewhere on the "structured-unstructured" spectrum. In the early stage of RPA adoption, structured audit tasks are better candidates[12]. Therefore, reconciliations, internal control testing, and detail testing can be appropriate targets for RPA (Moffitt et al. 2018).

Furthermore, auditors may consider extending the scale of "RPA-enabled" audit procedures and collect more evidence (such as from sampling to full population testing),

[12] Accounting firms may consider applying RPA to semi-structured tasks as well, when they have learned sufficient knowledge in the early stage of adoption.

because the limited processing power of human auditors is no longer a constraining factor. Additionally, the audit firm may examine the current audit process and reengineer the audit program to match the RPA software. For instance, when implementing the CA software at Siemens, Alles et al. (2008) suggest that the modification of the current audit process is necessary for the large-scale implementation of commercial CA software.

Stage 2: Data understanding

After selecting appropriate audit procedures, the audit firm should examine another important component of RPA-based audits: whether the data used in those procedures is compatible with RPA software. Data should be in a digital format or be able to be efficiently transformed into the digital content (Moffitt et al. 2018). Even though RPA commercial software is able to extract and interpret information from unstructured sources (such as images), structured data is still needed for accuracy and to minimize processing cost (Vinutha 2017). Therefore, although a procedure is automatable, it may not be feasible for RPA if the data are not in a digital format or are unstructured because such data may cause a high error rate and processing costs. In addition to the digital format, standardization of the data is essential. Moffitt et al. (2018) suggest that labels may be inconsistent across objects since auditors collect data from multiple sources. For instance, the vendor name in the client's system may be Amazon, but the official name on the contract may be Amazon.com, Inc. While human auditors can easily understand that Amazon and Amazon.com, Inc. are the same company, the RPA software may not be able to do so. Therefore, accounting firms need to check data consistency before the implementation of RPA.

Stage 3: Implementation

The implementation of RPA on selected audit procedures can be in-house or outsourced to an RPA provider. RPA service providers, including IBM, Genpact, Uipath, and Blue Prism, offer “business process-as-a-service” or “robots-as-a-service” arrangements (Santos 2017). Audit firms that want to implement RPA in-house may use free RPA software (such as Uipath Community Edition) or purchase licenses from RPA vendors. Usually, RPA software includes a user-friendly interface with a recording function that generates a script as a user performs the task that is to be automated (Moffitt et al. 2018). The main benefits of in-house implementation are that the accounting firm has a high level of control and that confidential information is better protected, but additional training for RPA designers is needed (Lacity et al. 2015). Outsourcing the RPA implementation to a third party can be faster but the costs may be higher. Table 18 displays the pros and cons of outsourcing and in-house RPA implementation summarized by practitioners (Santos 2017).

Table 18 Cons and Pros of Outsourcing and In-House RPA Implementation

	Pros	Cons
In-house implementation	• Changes in implementation can be done quickly and easily • The software robotics will be tailored for the business. • Easier quality control • Confidential proprietary information is protected • Freedom from being financially tied to another company • Quick reintegration of new tasks	• Additional financial investment for training employees • The need for additional workers • Adding training hours that may slow down work productivity
Outsourcing	• Fast reaction to changes in the scale of work • Fixed cost • More knowledgeable and experienced in the robotic technology industry • Retraining own employees not needed • Replaceable when needs aren't met	• Higher implementation cost • Additional expenses in contracts and other legal matters • Extra costs for robotic setup changes

Audit firms should be aware that even though RPA is relatively easy to implement, the implementation process involves risks and takes time. Many implementations actually fail, and those that succeed can take 4 to 6 months (PwC 2017; Srivastava 2017). Based on 18 RPA implementation projects by PwC (2017), the recommendation in this framework is that new adopters demonstrate the usability of RPA with a small part of a simple process through a proof-of-concept (PoC) or pilot project. During the start-up project, accounting firms will learn more about the RPA technology and software.

Stage 4: Evaluation and feedback

Finally, the audit teams need to evaluate the effectiveness of the RPA implementation (PoC or pilot project). Auditors may manually perform the procedures and compare the difference between their results and that of the software to determine the quality of the implantation. If the evaluation results reveal that the implementation needs improvement, the audit team needs to repeat steps 1 through 3—to modify audit procedures, change the data format/consistency, and adjust the implementation process—until the RPA software performs the procedure as expected.

5.5. The Evaluation and the Pilot Project

To confirm its effectiveness, the framework was evaluated by researchers and auditors and revised based on their feedback and comments. In addition, guided by the framework, we worked with a CPA firm to implement a pilot RPA project to perform the confirmation process.

Stage1: Procedure selection

We first observed the audit engagement of one client in the retail industry. Based on our observation and discussion with the audit team, we selected two audit procedures that match the criteria for candidates for RPA implementation: well-defined, repetitive, and mature.

The first procedure is the confirmation process, which confirms that the bank account balances directly with a third-party intermediary. In general, evidence obtained from an independent source, such as a bank or a third-party, is considered more reliable than evidence obtained from an internal source at the audit client (PCAOB 2010). Based on the CPA firm's audit plan, this procedure contained the following steps: 1) prepare request form, 2) initiate confirmation requests through Confirmatio.com[13] based on the information provided by the request form, 3) wait for the confirmations, and 4) download documents and extract the account balance for further audit tests. This audit procedure is well-defined because all of the steps (sending requests, monitoring the results, downloading the confirmation, and extracting the balance) do not require professional judgment and can be performed based on explicit rules. Additionally, 14 cash accounts need to be confirmed for this client, and this CPA firm performs 100% cash account confirmation for all clients. It performs 500 to 750 cash account confirmations every year, which makes the procedure highly repetitive. Also, the output of this procedure is generally predictable because this process should not have errors once the request form is correct, and auditors only need to download the document and extract the account balance from the confirmation.

The second candidate was the inventory cut-off test. In this test, the audit team checks whether the receiving date on the client's system is accurate. Specifically, auditors need to 1) extract the receiving date of each item from the client's inventory system, 2) retrieve the delivery date by searching for the tracking number on the carrier's website, and 3) compare these two dates across systems to determine the accuracy of the receiving date. All the steps in this procedure can be performed based on precise rules and no complex judgment is needed. Furthermore, auditors do not expect many unprecedented situations, which suggests that this procedure is quite mature.

---

[13] Confirmation.com is a web-based audit confirmation solution that is relied on by more 10,000 accounting firms in more than 100 countries, bringing efficiency and security to the confirmation process for cash, debt, accounts receivable, and more than 40 other confirmation types (Hanes et al. 2014, p356)

As the framework suggests, auditors need to consider whether to modify the current procedure to better match the software or change the scale of the procedure to collect more audit evidence. For the first procedure, we decided to modify the request template based on the input of the electronic confirmation platform and change the Word file template to an Excel file to help the RPA program better access the information. In addition, three columns (Status, Confirm, and Balance) were added into the template. The RPA program will update the status of each request case in these three columns to help the auditor monitor the process. Figure 3 shows the original and redesigned request templates.

Original request form

Company Information:

SIMPLY SOUPS INC.

177 WASHINGTON LANE

CHERRY HILL, NJ 08034

JENNINGS LOU

MANAGER

lou.jennings@ssoups.com

Phone: 609-555-5555

Bank Information:

FIFTH FEDERAL

Checking Account Number:

675-42223

Date for Confirm

(12/31/2016)

Re-designed request form

Client	Simply Soups Inc.
Signer Job Title	Manager
First Name	Jennings
Last Name	Lou
Address	177 Washington Lane
City	Cherry Hill
State	NJ
Zip code	8034
Email	lou.jennings@ssoups.com
Phone	609-555-5555
Bank Name	Fifth Federal
Form	Asset
Account ID	675-42223
Date	12/31/2016
Status	
Confirm	
Balance	

Figure 3 Original and Re-designed Request Form

For the second procedure, the original audit plan required auditors to select only 20 items to perform the cut-off test. Automation could increase the limited processing power of the auditors and they could collect more comprehensive audit evidence by increasing the scale of the test from sampling to full population testing.

Stage 2: Data understanding

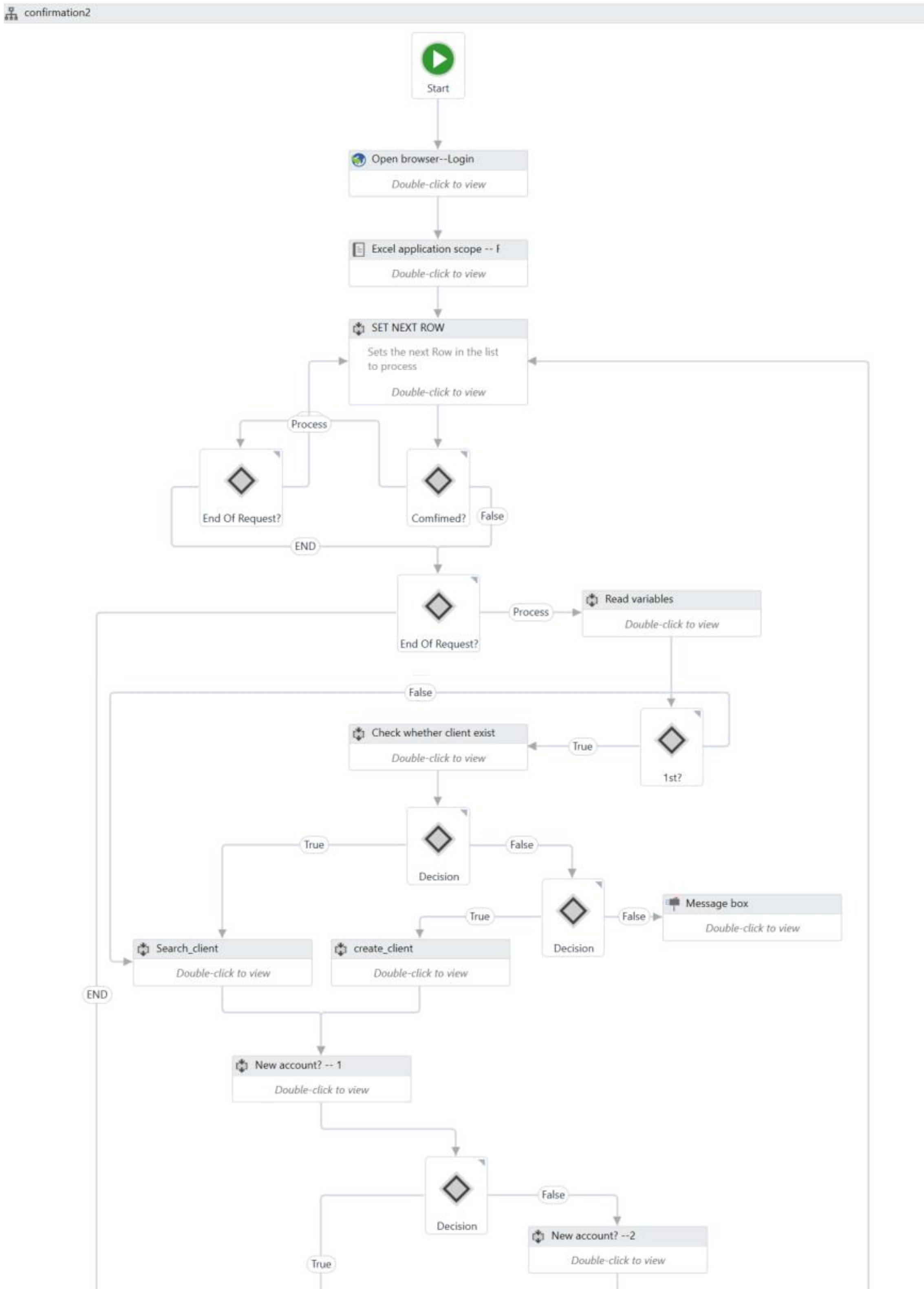
After selecting the procedures, the research team and the audit team examined the data used in each procedure. For the confirmation process, the data (client information, bank information, and account information) are in digital format in the request form, which is an Excel sheet. Although not all the information in the Confirmation.com webpage is in a structured format, most RPA software, equipped with optical character recognition (OCR), is able to transfer unstructured information (such as images) to textual format. Once the confirmation is downloaded, the account balance can be extracted from the PDF confirmation. Therefore, the data used in the confirmation process are all in digital format and can be handled by RPA software.

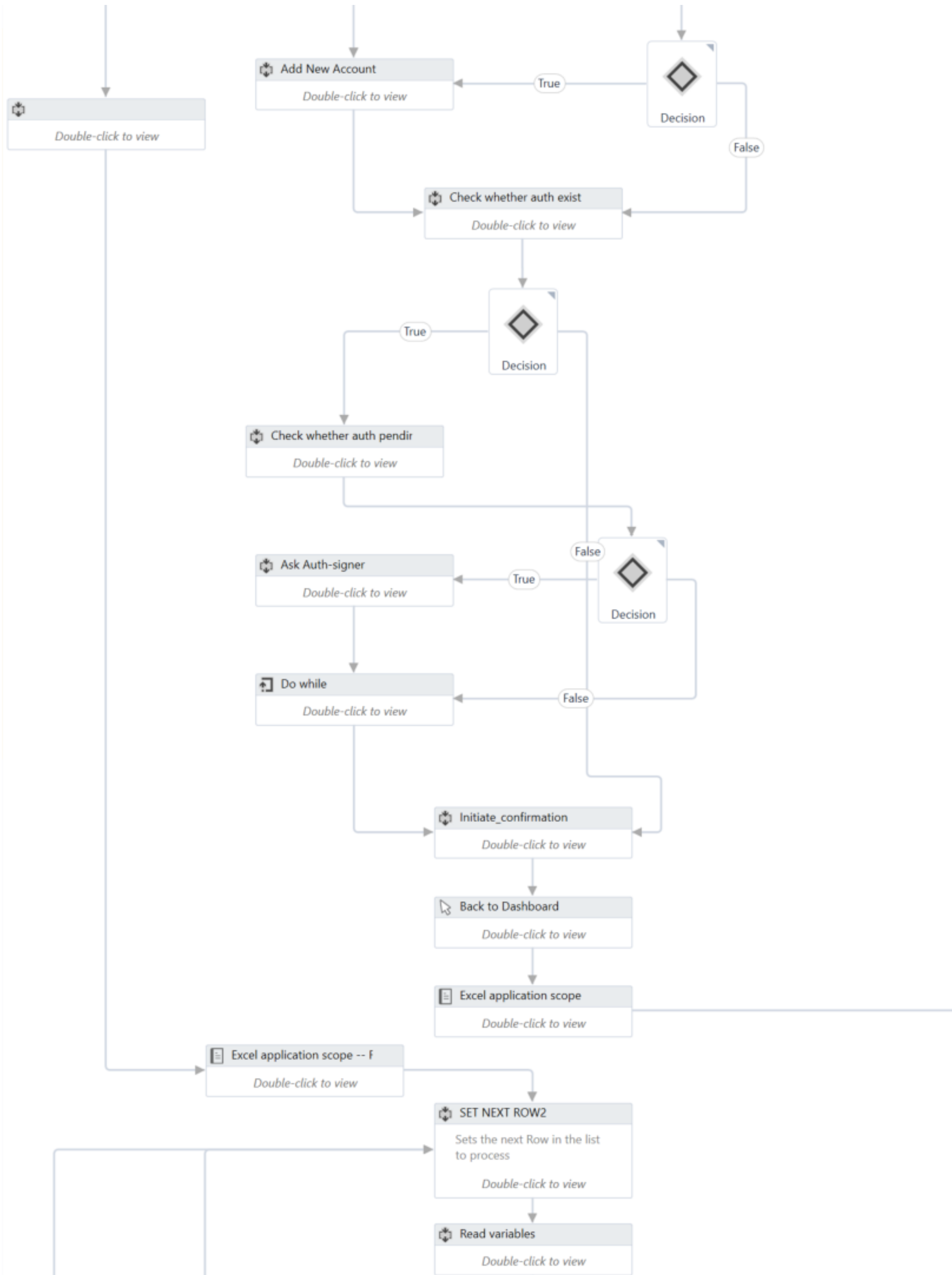
However, the data used in the second procedure, the inventory cut-off test, may not currently be feasible for RPA software. It's the inventory system for this client is still paper-based. If auditors want to automate this procedure, all the paper-based documents need to be transferred into digital format, which may not be cost-effective. Additionally, the tracking numbers, which are used to retrieve the delivery history on the carrier's website, are not available for all receiving items. Therefore, the data for this procedure leads us to decide that the inventory cut-off test is not a good candidate for the pilot implementation.

Finally, considering the characteristics of audit procedures and the data used in each process, the research team and the audit team agreed to implement RPA on the confirmation process as the pilot project.

Stage 3: Implementation

The implementation of this pilot project was in-house. We first selected a popular RPA software package, UiPath Community Edition, which is free for individual developers, small professional teams, and education entities. Then, we split the entire procedure into small pieces, and prepared a flow chart. Specifically, this procedure is divided into the following nine small steps: 1) open web browser and log in to confirmation.com, 2) extract the information from the pre-prepared request form, 3) check whether the client portfolio exists (if not, generate the client portfolio), 4) check whether the bank account exists (if not, add a new account), 5) check whether client authorization has been granted (if not, request authorization from the client), 6) initiate confirmation, 7) monitor the pending requests, 8) download the completed confirmations, and 9) extract account balances. Figure 4 presents this flowchart in detail.


confirmation2
Start
Open browser--Login
Double-click to view
Excel application scope -- F
Double-click to view
SET NEXT ROW
Sets the next Row in the list to process
Double-click to view
Process
End Of Request?
Comfimed?
False
END
End Of Request?
Process
Read variables
Double-click to view
False
Check whether client exist
Double-click to view
True
1st?
True
Decision
False
True
Decision
False
Message box
Double-click to view
Search_client
Double-click to view
create_client
Double-click to view
END
New account? -- 1
Double-click to view
Decision
False
New account? --2
Double-click to view
True



Add New Account
Double-click to view
True
Decision
False
Double-click to view
Check whether auth exist
Double-click to view
True
Decision
Check whether auth pendir
Double-click to view
False
Ask Auth-signer
Double-click to view
True
Decision
Do while
Double-click to view
False
Initiate_confirmation
Double-click to view
Back to Dashboard
Double-click to view
Excel application scope
Double-click to view
Excel application scope -- F
Double-click to view
SET NEXT ROW2
Sets the next Row in the list to process
Double-click to view
Read variables
Double-click to view

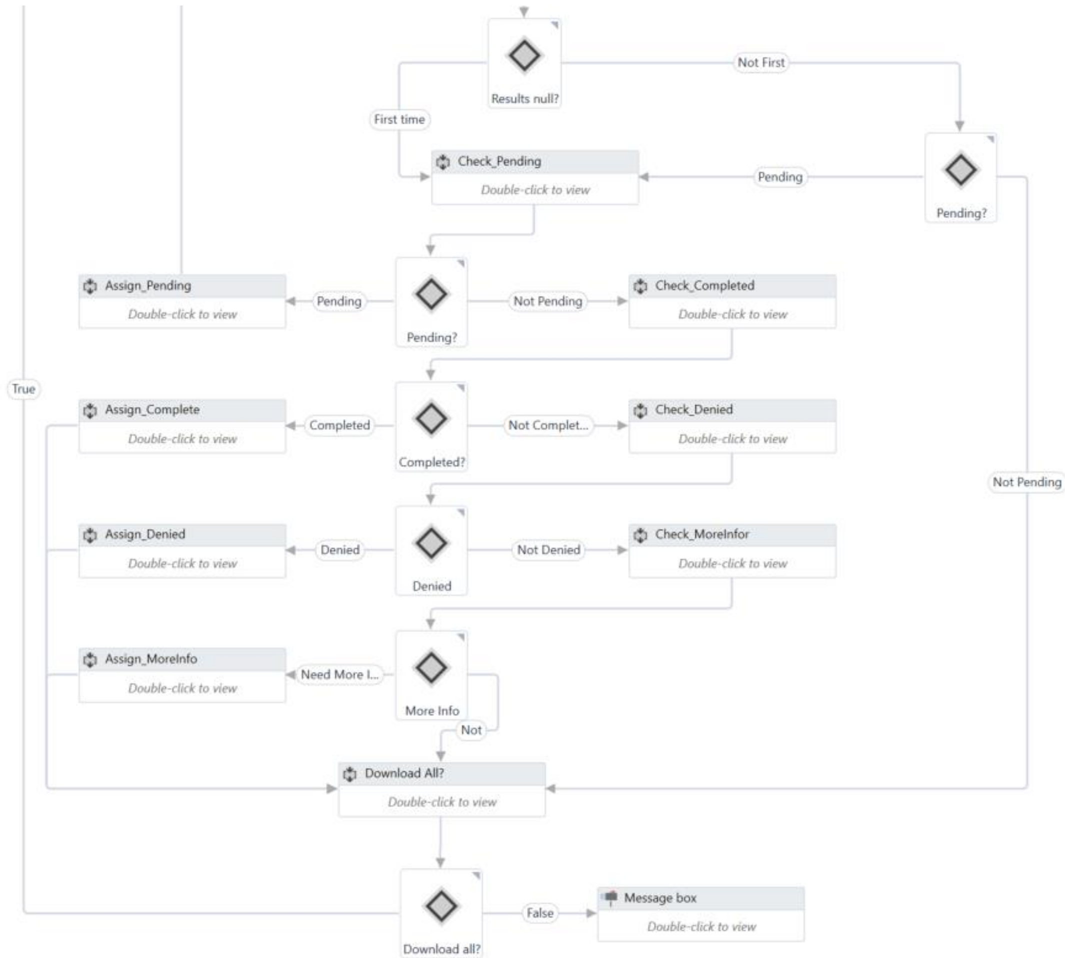


Figure 4 Flowchart for RPA Implementation of Confirmation Process

For each step, we chose appropriate activities in UiPath to mimic the way humans perform. For instance, to perform the first step "open browser and login Confirmation.com," we directed the software to conduct the following activities: 1) open Google Chrome browser, 2) go to URL: www.edu.confimation.com[14], 3) type userID, 4) type password, and 5) click "Login" button. Figure 5 displays the activities for this step.

[14] In this pilot project, we use edu.confirmation.com, which is an educational version of confirmation.com.

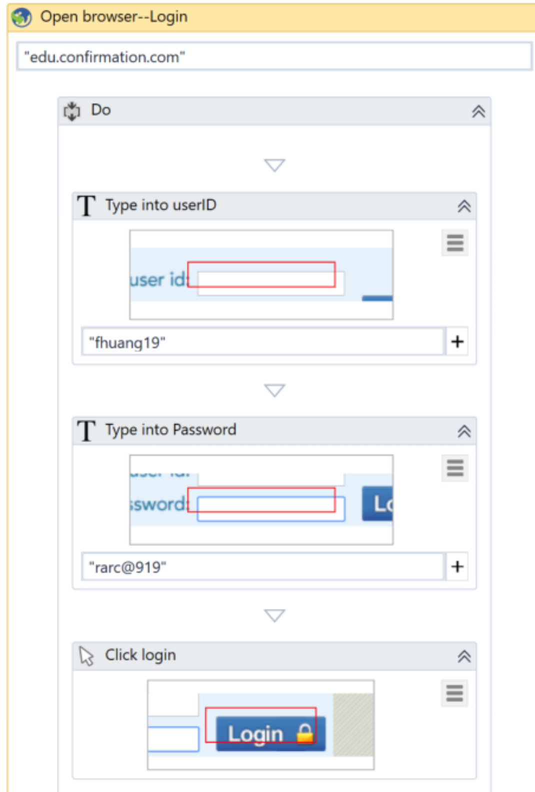


Figure 5 The Activities in the "Open browser and login" Step

For some steps, it is necessary to make a judgment, but these judgments can be made based on explicit rules. For example, to check whether the client portfolio exists, the program needs to search for the client name and read the output using the OCR. If the output shows that no existing record for this client exists, the program will automatically trigger the activity of generating a new client portfolio.

The pilot RPA program that was implemented to automatically perform the confirmation process is able to handle multiple confirmation requests. Once a confirmation request is initiated, the RPA program starts to extract the next available request from the request form, and repeats the request process. After all the requests have been submitted, the monitoring process checks the status of the confirmations until all of them have been completed and downloaded. Finally, the account balance is extracted to the request form and the status is updated to "Completed." Auditors only need to check the request form and follow up on any uncompleted confirmations, which will be labeled as "denied" or "more information needed."

Stage 4: Evaluation

The evaluation process for this RPA pilot project was based on a teaching case designed by Hanes et al. (2014) to perform a financial statement audit for Simply Soups, Inc., an American producer of organic canned soups. In this case, auditors need to follow the PCAOB's proposed standard on confirmations to complete the testing of the cash balances reported by Simply Soups, Inc. at year-end. They are required to complete the testing of Simply Soups, Inc.'s six bank accounts as of December 31, 2016, by using edu.confirmation.com to initiate confirmation and then to evaluate of responses received. The Appendix C shows the account information[15].

The request form based on the client's accounts information was prepared and the RPA program was initiated. Consistent with the expectation, the pilot program logged into the electronic confirmation platform – edu.confirmation.com, send confirmation requests, extracted the received account balances, and downloaded the completed confirmations. Then, the effectiveness of this pilot program was evaluated. Specifically, we performed the entire process manually and compared the received documents with the ones collected from the RPA program. The results show that all the information from the completed cases (such as client name, bank names, account numbers, date, account balances) match, supporting the accuracy of this RPA pilot project.

[15] The audit year end date in the teaching case, which is based on Hanes et al. (2014), has been updated by edu.confirmation.com to December 31, 2016.

During the evaluation of the pilot project, we encountered two exceptions. First, one bank was no longer included in the network of edu.confirmation.com, so auditors needed to deal with this issue manually and mail a paper-based confirmation request. The second exception was that the information about the authorized signer was incorrect, which is not uncommon for the confirmation process because clients sometimes forget to update the information to auditors. In that situation, auditors need to contact the client for the correct information[16]. Overall, the evaluation of this pilot RPA project demonstrated the feasibility of RPA in the auditing area. The audit team could extend the scale of this pilot RPA project to all confirmations across different clients and save auditors' time for tasks requiring professional judgment.

5.6. Discussion and Conclusion

5.6.1 The Benefits of RPA in Auditing

With the application of RPA, software will automatically perform the pre-designed audit tasks, and audit procedures will no longer be limited by human processing power. Currently, because of limited audit resources, sampling techniques (statistical sampling and non-statistical sampling) are commonly used in audit processes such as tests of controls and substantive tests of details (Christensen et al. 2014). With the adoption of RPA, auditors can expand the scale of some procedures from sampling to testing the entire population, avoiding the risks and deficiencies related to sampling and collecting comprehensive audit evidence.

Another benefit of RPA is to avoid human errors such as mistakes when confirming

[16] A discussion of how to deal with exceptions or errors is provided in the next section.

amounts, errors in workpapers, and ignoring red flags. Once the RPA software is programmed precisely, the tasks and analyses will be performed in line with audit standards and pre-defined rules, helping to mitigate the occurrence of human errors. recent survey from major accounting firms shows that, compared with the 90 percent accuracy rate of humans, RPA accuracy achieves 99.9 percent (Cooper et al. 2018). Finally, RPA offers scalability and flexibility, which allow accounting firms to re-assign RPA software for different analyses or different audit engagements without significant costs.

5.6.2 Errors and Exceptions

Even though the ideal candidates for RPA are structured and well-defined, it is difficult to avoid all errors or exceptions during the application. The first solution is to automate the normal and routine process and leave all the errors and exceptions to human auditors if the errors or exceptions are encountered infrequently. For example, during the confirmation process, it is possible to receive feedback that the requests cannot be completed because the bank no longer exists in the network. In this case, the RPA software only identifies the uncompleted requests and leaves the follow-up for auditors. Specifically, the type of error (such as request denied, more information needed) and the comments (such as invalid date or invalid contactor) are sent to auditors, and all the follow-up is performed manually.

However, if the errors or exceptions happen regularly, leaving them for the manual process may not be efficient. The solution for regularly occurring errors and exceptions is to first classify all errors/exceptions into two groups: common errors and uncommon errors. Then, the follow-up procedures for the common errors should be included in the RPA automation. Once this type of error is detected, the follow-up process will be triggered automatically. Then, auditors need to deal only with uncommon errors, and the efficiency of the RPA application is enhanced.

5.6.3 RPA’s Next Step - IPA

Although RPA is potentially making significant improvements in auditing practice, a key limitation currently is that the software is able to perform only routine tasks and to make decisions based on explicit rules. Therefore, current RPA software is not adaptable to audit procedures requiring professional judgment because that judgment cannot be transformed into structured instructions.

Recently, there has been a progress in the evolution of technology aimed at applying artificial intelligence (AI) in the industry. Large accounting firms have launched numerous projects to implement AI in their audit practice (Kokina and Davenport 2017). For instance, KPMG is working with IBM Watson to apply cognitive computing technology to its professional services offerings (IBM 2016). Deloitte is collaborating with Kira Systems, a contract analysis system, to create cognitive models that examine large numbers of complex documents, extract textual information for better analysis, and assist auditors with the difficult task of document reviewing (Deloitte 2016). To take advantage of AI developments and address the limitations of current RPA, practitioners have proposed Intelligent Process Automation (IPA), which refers to the combination of AI, cognitive automation, deep learning, and machine learning with RPA (e.g., Berruti et al. 2017; UiPath 2018).

Instead of only mimicking the way people perform routine business processes, IPA leverages the advantages of AI to learn how people make decisions and may be able to perform complex tasks faster and better. To further improve audit quality, accounting firms may consider applying IPA in the future to help auditors perform the complex and unstructured audit procedures and make professional judgments.

5.6.4 Concluding Remarks

This paper introduces an emerging automation technology, robotic process automation (RPA), to the auditing practice. RPA is a method that performs routine business processes by automating the way that people interact with multiple applications or systems through a user interface and also by following simple rules to make decisions (Deloitte 2017). Although the benefits of RPA have been documented in different industries and many audit tasks (such as reconciliations, internal control testing, and detail testing) are ideal candidates for RPA, applications of RPA in auditing remain unexplored.

This study proposes a framework to guide auditors in applying RPA. Specifically, firms need to 1) select well-defined, repetitive, and mature audit procedures, 2) understand the data used in each procedure, 3) implement a pilot or PoC project for a simple process, and 4) evaluate the effectiveness and efficiency of the project. Finally, the feasibility of RPA in the audit practice is demonstrated by implementing a pilot RPA project to automate the confirmation process.

This essay is not without limitations. First, the pilot project was built to perform the confirmation process and can only demonstrate the effectiveness of this one audit procedure. Second, the RPA software used in the pilot project has a limited capability for dealing with errors. Future research could extend RPA implementation to other audit procedures across multiple audit stages. It could be important for future studies to design follow-up RPA-enabled audit procedures to minimize manual intervention from auditors. Finally, the long-term benefits of RPA implementation, which involves cost saving over multiple audit engagements and improved audit quality, could be interesting for future research.

CHAPTER 6: CONCLUSION

The field of accounting is undergoing a fundamental change due to the advances in emerging technologies. This dissertation studies the impact of emerging technologies in accounting and examines their effects on the financial reporting process and the audit practice.

The first two essays examine the effect of XBRL on managers' behavior. Specifically, my first essay focuses on a current debate in the literature and investigates whether managers use XBRL elements strategically to increase the complexity of the interactive data format of financial statements (i.e., XBRL filings). The SEC's XBRL mandate, which requires firms to tag each financial fact in their financial statements using either a standard element or an extension element, offers a unique setting for our research questions and allows us to examine managers' discretionary use of extension elements without a self-selection bias. Using the ratio of extension elements to total elements in an XBRL 10-K filing as the measure of XBRL complexity, this chapter finds evidence that firms' XBRL filings are more complex when the firms are performing poorly and when their good news (bad news) is less (more) persistent. The findings further show that the effect is more pronounced when firms are more inherently complex. The results suggest that managers tend to use extension elements strategically to introduce XBRL complexity and obfuscate XBRL-tagged financial information.

The second essay fills a research gap by utilizing the adoption of XBRL as an exogenous shock to examine the effect of information processing efficiency on firms' investment behavior. The results indicate that the adoption of XBRL can significantly reduce abnormal investments, especially over-investments. This chapter further investigates potential factors that may magnify or mitigate the benefits of XBRL adoption on investment efficiency. The findings show that such benefits will be mitigated in firms with superior external monitoring, less environment uncertainty, and more linguistically complex reporting. These results are robust to a "difference in difference" empirical setting and continue to hold for non-capital investments. Considering the learning ability of investors in understanding and interpreting XBRL, the results also show a persistent and monotonically improved effect of XBRL adoption, supporting the expectation that the effect of XBRL adoption on investment efficiency enhances as time goes. Overall, the results of this study contribute to the existing accounting literature by directly examining the effects of the efficiency of information processing on investment efficiency.

The third essay introduces an emerging automation technology, robotic process automation (RPA), to the auditing practice. RPA is a method that performs routine business processes by automating the way that people interact with multiple applications or systems through a user interface and also by following simple rules to make decisions (Deloitte 2017). A framework is proposed to apply RPA to automate repetitive and well-defined procedures in order to free auditors from doing repetitive and low-judgment audit tasks and enable them to focus on audit tasks that require professional judgment. Finally, this study demonstrates the feasibility of RPA by implementing a pilot project that applies RPA to the confirmation process.

In conclusion, this dissertation examines the effects of XBRL on financial reporting strategy and managers' investment decision, proposes to apply RPA to automate labor-intensive, well-defined and repetitive audit procedures, and demonstrate the feasibility of RPA in the audit practice.

BIBLIOGRAPHY

Abdolmohammadi, M. J. 1999. A comprehensive taxonomy of audit task structure, professional rank and decision aids for behavioral research. *Behavioral Research in Accounting* 11:51-92.

Aboody, D., and B. Lev. 2000. Information asymmetry, R&D, and insider gains. *The Journal of Finance* 55 (6):2747-2766.

Agoglia, C. P., J. F. Brazel, R. C. Hatfield, and S. B. Jackson. 2010. How do audit workpaper reviewers cope with the conflicting pressures of detecting misstatements and balancing client workloads? *Auditing: A Journal of Practice & Theory* 29 (2):27-43.

AICPA. 2007. XBRL Basics - How does it work? .

———. 2015. Audit analytics and continuous audit.

———. 2017. Audit Data Analytics Guide.

Ali, A., and W. Zhang. 2015. CEO tenure and earnings management. *Journal of Accounting and Economics* 59 (1):60-79.

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

Appelbaum, D., A. Kogan, M. Vasarhelyi, and Z. Yan. 2017. Impact of business analytics and enterprise systems on managerial accounting. *International Journal of Accounting Information Systems* 25:29-44.

Appelbaum, D., and R. A. Nehmer. 2017. Using drones in internal and external audits: An exploratory framework. *Journal of Emerging Technologies in Accounting* 14 (1):99-113.

Bae, G. S., S. U. Choi, D. S. Dhaliwal, and P. T. Lamoreaux. 2016. Auditors and client investment efficiency. *The Accounting Review* 92 (2):19-40.

Bagnoli, M., M. B. Clement, and S. G. Watts. 2005. Around-the-clock media coverage and the timing of earnings announcements. *Working paper*.

Ball, R., S. Jayaraman, and L. Shivakumar. 2012. Audited financial reporting and voluntary disclosure as complements: A test of the confirmation hypothesis. *Journal of Accounting and Economics* 53 (1):136-166.

Bedard, J. C., M. Ettredge, and K. M. Johnstone. 2006. Using electronic audit workpaper systems in audit practice: Task analysis, learning, and resistance. *Working paper*.

Berruti, F., G. Nixon, G. Taglioni, and R. Whiteman. 2017. *Intelligent process automation: The engine at the core of the next-generation operating model*. Available from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/intelligent-process-automation-the-engine-at-the-core-of-the-next-generation-operating-model.

Biddle, G. C., C. M. Callahan, H. A. Hong, and R. L. Knowles. 2011. Does mandatory adoption of international financial reporting standards increase investment efficiency. *Working paper*.

Biddle, G. C., and G. Hilary. 2006. Accounting quality and firm-level capital investment. *The Accounting Review* 81 (5):963-982.

Biddle, G. C., G. Hilary, and R. S. Verdi. 2009. How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics* 48 (2):112-131.

Blankespoor, E., B. P. Miller, and H. D. White. 2014. Initial evidence on the market impact of the XBRL mandate. *Review of Accounting Studies* 19 (4):1468-1503.

Blankley, A. I., D. N. Hurtt, and J. E. MacGregor. 2012. Abnormal audit fees and restatements. *Auditing: A Journal of Practice & Theory* 31 (1):79-96.

Bloomfield, R. J. 2002. The "incomplete revelation hypothesis" and financial reporting. *Accounting Horizons* 16 (3):233-243.

Boritz, J. E., and W. G. No. 2008. The SEC's XBRL voluntary filing program on EDGAR: A case for quality assurance. *Current Issues in Auditing* 2 (2):A36-A50.

———. 2009. Assurance on XBRL-related documents: The case of United Technologies Corporation. *Journal of Information Systems* 23 (2):49-78.

Boritz, J. E., and L. M. Timoshenko. 2014. Firm-specific characteristics of the participants in the SEC's XBRL voluntary filing program. *Journal of Information Systems* 29 (1):9-36.

Bozanic, Z., and M. Thevenot. 2015. Qualitative disclosure and changes in Sell-Side financial analysts' information environment. *Contemporary Accounting Research* 32 (4):1595-1616.

Burgess, A. 2016. Time to talk – RPA and AI in contact centers.

Bushman, R. M., and A. J. Smith. 2001. Financial accounting information and corporate governance. *Journal of Accounting and Economics* 32 (1):237-333.

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

Chang, X., S. Dasgupta, and G. Hilary. 2009. The effect of auditor quality on financing decisions. *The Accounting Review* 84 (4):1085-1117.

Chasan, E. 2013. Costly data go untapped. *The Wall Street Journal*.

Chen, F., O.-K. Hope, Q. Li, and X. Wang. 2011. Financial reporting quality and investment efficiency of private firms in emerging markets. *The Accounting Review* 86 (4):1255-1288.

Chen, J. Z., H. A. Hong, J.-B. Kim, and J. W. Ryou. 2016. Information processing costs and corporate tax aggressiveness: Evidence from the SEC's XBRL mandate. *Working paper*.

Cheng, M., D. Dhaliwal, and Y. Zhang. 2013. Does investment efficiency improve after the disclosure of material weaknesses in internal control over financial reporting? *Journal of Accounting and Economics* 56 (1):1-18.

Christensen, B. E., R. J. Elder, and S. M. Glover. 2014. Behind the numbers: Insights into large audit firm sampling policies. *Accounting Horizons* 29 (1):61-81.

CICA/AICPA. 1999. Continuous auditing.

Cong, Y., J. Hao, and L. Zou. 2014. The impact of XBRL reporting on market efficiency. *Journal of Information Systems* 28 (2):181-207.

Cooper, L., D. Holdernedd, T. Sorensen, and D. Wood. 2018. Robotic process automation in public accounting. *Working paper.*

Debreceny, R. S., S. M. Farewell, M. Piechocki, C. Felden, A. Gräning, and A. d'Eri. 2011. Flex or break? Extensions in XBRL disclosures to the SEC. *Accounting Horizons* 25 (4):631-657.

Dechow, P. M., and I. D. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (s-1):35-59.

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

Deloitte. 2016. Deloitte forms alliance with Kira Systems.

———. 2017. Automate this: The business leader’s guide to robotic and intelligent automation.

Desmond, L. 2014. XBRL health check: Extension tags. *CERTENT Disclosure Net*.

Dhole, S., G. J. Lobo, S. Mishra, and A. M. Pal. 2015. Effects of the SEC's XBRL mandate on financial reporting comparability. *International Journal of Accounting Information Systems* 19:29-44.

Dong, Y., O. Z. Li, Y. Lin, and C. Ni. 2016. Does information-processing cost affect firm-specific information acquisition? Evidence from XBRL adoption. *Journal of Financial and Quantitative Analysis* 51 (02):435-462.

Du, H., M. A. Vasarhelyi, and X. Zheng. 2013. XBRL mandate: Thousands of filing errors and so what? *Journal of Information Systems* 27 (1):61-78.

Efthimides, E. 2017. Q&A with an expert: Bloomberg Global Data's use of XBRL shows the vital importance of data quality.

EY. 2016a. Get ready for robots.

———. 2016b. Robotic process automation in the finance function of the future.

Feng, M., C. Li, and S. McVay. 2009. Internal control and management guidance. *Journal of Accounting and Economics* 48 (2):190-209.

Ferreira, M. A., and P. Matos. 2008. The colors of investors' money: The role of institutional investors around the world. *Journal of Financial Economics* 88 (3):499-533.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39 (2):295-327.

Francis, J., D. Nanda, and P. Olsson. 2008. Voluntary disclosure, earnings quality, and cost of capital. *Journal of Accounting Research* 46 (1):53-99.

Gomariz, M. F. C., and J. P. S. Ballesta. 2014. Financial reporting quality, debt maturity and investment efficiency. *Journal of Banking & Finance* 40:494-506.

Gregor, S., and A. R. Hevner. 2013. Positioning and presenting design science research for maximum impact. *MIS Quarterly* 37 (2):337-355.

Gunning, R. 1952. The technique of clear writing.

Guragai, B., N. Hunt, M. Neri, and E. Z. Taylor. 2014. Accounting information systems and ethics research: Review, synthesis, and the future. *Journal of Information Systems* 31 (2):65-81.

Hanes, D. R., B. M. Porco, and J. C. Thibodeau. 2014. Simply Soups Inc.: A teaching case designed to integrate the electronic cash confirmation process into the auditing curriculum. *Issues in Accounting Education* 29 (2):349-369.

Healy, P. M., and K. G. Palepu. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31 (1):405-440.

Hevner, A. R., S. T. March, J. Park, and S. Ram. 2004. Design science in information systems research. *MIS Quarterly* 28 (1):75-105.

Hodge, F. D., J. J. Kennedy, and L. A. Maines. 2004. Does search-facilitating technology improve the transparency of financial reporting? *The Accounting Review* 79 (3):687-703.

Hoitash, R., and U. Hoitash. 2017. Measuring accounting reporting complexity with XBRL. *The Accounting Review* 93 (1):259-287.

Holmstrom, B. 1982. Moral hazard in teams. *The Bell Journal of Economics*:324-340.

Hope, O. K., and W. B. Thomas. 2008. Managerial empire building and firm disclosure. *Journal of Accounting Research* 46 (3):591-626.

Hou, Q., Q. Jin, L. Wang, and G. Zhang. 2016. Mandatory IFRS adoption, accounting quality, and investment efficiency: Evidence from China. *China Journal of Accounting Studies* 4 (3):236-262.

Hubbard, R. G. 1997. Capital-market imperfections and investment: National Bureau of economic research.

IBM. 2016. KPMG announces agreement with IBM Watson to help deliver cognitive-powered insights.

IEEE Corporate Advisory Group. 2017. *IEEE guide for terms and concepts in intelligent process automation*. New York City: IEEE.

IRPA. 2015. Introduction to robotic process automation: A primer.

———. 2016. What is Robotic Process Automation?

Issa, H., T. Sun, and M. A. Vasarhelyi. 2016. Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation. *Journal of Emerging Technologies in Accounting* 13 (2):1-20.

Jensen, M. C. 1986. Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review* 76 (2):323-329.

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

Jung, B., W. J. Lee, and D. P. Weber. 2014. Financial reporting quality and labor investment efficiency. *Contemporary Accounting Research* 31 (4):1047-1076.

Jung, M. J., J. P. Naughton, A. Tahoun, and C. Wang. 2016. Do firms strategically disseminate? Evidence from corporate use of social media. *Working paper*.

Kasznik, R. 1999. On the association between voluntary disclosure and earnings management. *Journal of Accounting Research* 37 (1):57-81.

Khurana, I. K., and W. J. Moser. 2012. Institutional shareholders' investment horizons and tax avoidance. *The Journal of the American Taxation Association* 35 (1):111-134.

Kim, J.-B., J. W. Kim, and J.-H. Lim. 2013. Does XBRL adoption constrain managerial opportunism in financial reporting? Evidence from mandated US filers. *Working paper*.

Kim, J. W., J.-H. Lim, and W. G. No. 2012. The effect of first wave mandatory XBRL reporting across the financial information environment. *Journal of Information Systems* 26 (1):127-153.

Kirk, M., J. Vincent, and D. Williams. 2016. From print to practice: XBRL extension use and analyst forecast properties. *Working paper*.

Kokina, J., and T. H. Davenport. 2017. The emergence of artificial intelligence: How automation is changing auditing. *Journal of Emerging Technologies in Accounting* 14 (1):115-122.

Kothari, S. P., A. J. Leone, and C. E. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1):163-197.

KPMG. 2017. KPMG and Automation Anywhere form alliance to deliver robotic process automation solutions to clients.

Lacity, M., L. P. Willcocks, and A. Craig. 2015. Robotic process automation at Telefonica O2.

Lambert, R., C. Leuz, and R. E. Verrecchia. 2007. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45 (2):385-420.

Lara, J. M. G., B. G. Osma, and F. Penalva. 2016. Accounting conservatism and firm investment efficiency. *Journal of Accounting and Economics* 61 (1):221-238.

Lawrence, A. 2013. Individual investors and financial disclosure. *Journal of Accounting and Economics* 56 (1):130-147.

Lehavy, R., F. Li, and K. Merkley. 2011. The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review* 86 (3):1087-1115.

Lenger, S., J. Ernstberger, and J. Stiebal. 2011. The impact of IFRS on firms' investment efficiency-European evidence on private and public firms. *Illinois International Journal of Accounting Symposium*.

Lennox, C. S., J. R. Francis, and Z. Wang. 2011. Selection models in accounting research. *The Accounting Review* 87 (2):589-616.

Li, E. X., and K. Ramesh. 2009. Market reaction surrounding the filing of periodic SEC reports. *The Accounting Review* 84 (4):1171-1208.

Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45 (2):221-247.

Lim, C. Y., T. Y. Thong, and D. K. Ding. 2008. Firm diversification and earnings management: evidence from seasoned equity offerings. *Review of Quantitative Finance and Accounting* 30 (1):69-92.

Liu, C., T. Wang, and L. J. Yao. 2014. XBRL's impact on analyst forecast behavior: An empirical study. *Journal of Accounting and Public Policy* 33 (1):69-82.

Loughran, T., and B. McDonald. 2014. Measuring readability in financial disclosures. *The Journal of Finance* 69 (4):1643-1671.

Lundholm, R. J., R. Rogo, and J. L. Zhang. 2014. Restoring the tower of Babel: How foreign firms communicate with US investors. *The Accounting Review* 89 (4):1453-1485.

McCafferty, J. 2013. If you build it, they may not come. *Compliance Week*.

McNichols, M. F., and S. R. Stubben. 2008. Does earnings management affect firms' investment decisions? *The Accounting Review* 83 (6):1571-1603.

Moffitt, K., A. Rozario, and M. Vasarhelyi. 2018. Robotic process automation for auditing. *Journal of Emerging Technologies in Accounting* 15 (1):1-10.

Moffitt, K. C., and M. A. Vasarhelyi. 2013. AIS in an age of Big Data. *Journal of Information Systems* 27 (2):1-19.

No, W. G., K. Lee, F. Huang, and Q. Li. 2018. Multidimensional audit data selection (MADS): A framework for using data analytics in the audit data selection process. *Working paper*.

PCAOB. 2010. Audit Evidence No.15. Washington, DC: PCAOB.

Premuroso, R. F., and S. Bhattacharya. 2008. Do early and voluntary filers of financial information in XBRL format signal superior corporate governance and operating performance? *International Journal of Accounting Information Systems* 9 (1):1-20.

PwC. 2017. Successful implementation of RPA takes time.

Richardson, S. 2006. Over-investment of free cash flow. *Review of Accounting Studies* 11 (2-3):159-189.

Santos, M. 2017. *The Pros and Cons of In-House Versus Outsourced RPA*. Available from https://www.ideyatech.com/pros-cons-in-house-versus-outsourced-rpa/.

Scherr, E., and D. Ditter. 2017. Customization versus standardization in electronic financial reporting: Early evidence from the SEC XBRL mandate. *Journal of Information Systems* 32 (2):125-148.

Schrand, C. M., and B. R. Walther. 2000. Strategic benchmarks in earnings announcements: The selective disclosure of prior-period earnings components. *The Accounting Review* 75 (2):151-177.

SEC. 2005. XBRL Voluntary Financial Reporting Program on the EDGAR System. Washington, DC: SEC.

———. 2009. Interactive Data to Improve Financial Reporting. Washington, DC: SEC.

———. 2010. Staff Observations From Review of Interactive Data Financial Statements. Washington, DC: SEC.

———. 2011. Staff Observations From Review of Interactive Data Financial Statements. Washington, DC: SEC.

———. 2014. Staff Observations of Customer Tag Rates. Washington, DC: SEC.

Segal, B., and D. Segal. 2016. Are managers strategic in reporting non-earnings news? Evidence on timing and news bundling. *Review of Accounting Studies* 21 (4):1203-1244.

Simon, H. A. 1996. *The sciences of the artificial*: MIT press.

Skinner, D. J., and E. Soltes. 2011. What do dividends tell us about earnings quality? *Review of Accounting Studies* 16 (1):1-28.

Sloan, R. G. 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? *The Accounting Review* 71 (3):289-315.

Srinivasan, V. 2016. *The intelligent enterprise in the era of big data*: John Wiley & Sons.

Srivastava, S. 2017. Why RPA implementations fail. *CIO Magazine*.

Stein, J. C. 2003. Agency, information and corporate investment. *Handbook of the Economics of Finance* 1:111-165.

UiPath. 2018. *Intelligent Process Automation -- Automation enabled by context-aware robots*. Available from https://www.uipath.com/automate/intelligent-process-automation.

Vasarhelyi, M. A. 1984. Automation and changes in the audit process. *Auditing: A Journal of Practice & Theory* 4 (1):100-106.

Vasarhelyi, M. A., M. G. Alles, and A. Kogan. 2004. Principles of analytic monitoring for continuous assurance. *Journal of Emerging Technologies in Accounting* 1 (1):1-21.

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

Vinutha. 2017. *Image Recognition – The heart of sophisticated RPA*. Available from https://blog.inatrix.com/image-recognition-rpa/.

Werbel, J., and D. B. Balkin. 2010. Are human resource practices linked to employee misconduct?: A rational choice perspective. *Human Resource Management Review* 20 (4):317-326.

XBRL.US. 2009. XBRL for filers: Implementing XBRL for reporting.

———. 2010. Avoiding common errors in XBRL creation.

———. 2014. The basics of XBRL for business and accounting professionals.

Yan, Z., K. Moffitt, and M. Vasarhelyi. 2017. Automated contract analysis in auditing. *Working paper*.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22:59-82.

APPENDICES

Appendix A Variable Definitions for Chapter 3

Variables		Definition
XBRL Complexity	*Exratio*	The number of extension elements divided by the total number of elements in an XBRL 10-K filing
Firm's Characteristics	*Earnings*	The operating income scaled by the book value of assets
	LnMVE	The natural logarithm of market value
	MTB	The market-to-book ratio.
	EarnVol	The standard deviation of quarterly ROA over the prior 7 years
	SI	The absolute value of special items scaled by total assets
	Merger	An indicator variable that equals one if the firm undertook a large merger or acquisition, and zero otherwise
	AbsAcc	The absolute amount of operating accruals
	Div	An indicator variable that equals one if the firm has a dividend this year, and zero otherwise
External Monitoring	*NAnalyst*	The natural logarithm of analysts following
XBRL Experience	*XBRLage*	The number of years between the current year and the year that the firm first files an XBRL-tagged 10-K
Complexity	*NBSeg*	The number of business segments
	NGSeg	The number of geographic segments
	Foreign	An indicator variable that equals one if the firm has foreign operations, and zero otherwise
	OpeComplex	A factor comprising the number of business segments, the number of geographic segments, and the existence of foreign business
	FogIndex	The Gunning Fog Index calculated as (words per sentence + percent of complex words) * 0.4
	Grosize	The natural logarithm of the file size of the EDGAR "complete submission text file" for the 10-K file
	Length	The natural logarithm of the number of words in the 10-K file
	LinComplex	A factor comprising the gross size and the natural logarithm of the number of words in the 10-K file

Appendix B Variable Definitions for Chapter 4

Variables	Definition
AbsXINV	The absolute value of residual value from investment model, following McNichols and Stubben (2008) and Bae et al. (2016).
XINV	The residual value from investment model, following McNichols and Stubben (2008) and Bae et al. (2016).
AbsNXINV	The absolute value of the residual value of non-capital investment from investment model.
Post	An indicator variable that equals 1 if firm-year is after firm's XBRL adoption and 0 otherwise.
Post1	An indicator variable that equals 1 if firm-year is one year after firm's XBRL adoption and 0 otherwise.
Post2	An indicator variable that equals 1 if firm-year is two years after firm's XBRL adoption and 0 otherwise.
Post3	An indicator variable that equals 1 if firm-year is three years after firm's XBRL adoption and 0 otherwise.
LnMve	The natural logarithm of the market value.
Loss	An indicator variable that equals 1 if net income before the extraordinary item is negative and 0 otherwise.
Leverage	Total liabilities scaled by total assets.
Cash	The ratio of cash to total assets.
MTB	The ratio of market value to book value.
Analyst	The number of analysts following.
StdCFO	The standard deviation of cash flows from operations scaled by average total assets from year t-5 to t-1.
StdSales	The standard deviation of sales scaled by average total assets from year t-5 to t-1.
StdInvestment	The standard deviation of investments from year t-5 to t-1.
Z-Score	The distress score developed by Zmijewski (1984).
Tangibility	The ratio of property, plant, and equipment to total assets.
OperCycle	The nature logarithm of receivables to sales plus inventory to COGS multiplied by 360.
FRQ	Accruals quality defined by Francis et al. (2005).
NBus	The number of business segments.
FogIndex	The Gunning Fog Index of 10-K file.
Length	The natural logarithm of the total number of words in the 10-K file.
InstHold	The percentage of a firm's share held by institutional investors.
Ferror	The analyst forecast errors of annual EPS.

Appendix C Key Account Information

Audit Firm:	Putnam and Jacobs LLP	
Year End Date:	December 31, 2016	
Client Address:	177 Washington Lane	Cherry Hill, NJ 08034
Client Phone Number:	609-555-5555	
Client Contacts:	Lou Jennings	Lou.jennings@ssoups.com
	Chuck Rogers	Chuck.rogers@ssoups.com

Bank Accounts – 2016

Bank Name	Bank Address	Bank Manager	Account Name	Account Number(s)	Authorized Signer
Fifth Federal	73 Union Street New York NY 10001	George Williams	Checking	675-42223	Lou Jennings
Sparkasse-Frankfurt	Landstrasse 89-21 Frankfurt 60326 DE	Helga Jones	Checking	44-322711	Lou Jennings
American NorthWest Bank	234 Market Street Milwaukee WI 53202	Richard Johnson	Checking	05-198305	Lou Jennings
BNY Federal	3621 Ave De Lafayette Boston MA 02111	Betty Smith	Savings	061-22031	Lou Jennings
Tenth National Bank	313 S. Chadwick Street Philadelphia PA 19103	Greg Fordham	Savings	26-798422	Lou Jennings
Bank of Citizens	3621 Union Ave Denver CO 80220	Denise Bentley	Checking	89-123661	Lou Jennings