AUDIT FOCUSED PROCESS MINING: THE EVOLUTION OF PROCESS MINING AND INTERNAL CONTROL

By

ABDULRAHMAN ALREFAI

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

Majoring in Accounting Information Systems

Written under the direction of

Professor Miklos A. Vasarhelyi

and approved by

_____

Professor Miklos A. Vasarhelyi

_____

Professor Helen Brown-Liburd

_____

Professor Hussein Issa

_____

Professor Rajendra P. Srivastava

Newark, New Jersey

May, 2019

ABSTRACT OF THE DISSERTATION

Audit Focused Process Mining: The Evolution of Process Mining and Internal Control

By Abdulrahman Alrefai

Dissertation Chairman: Professor Miklos A. Vasarhelyi

Process mining has been introduced as an auditing tool to aid auditors in examining the business processes effectively and efficiently. This dissertation will demonstrate in three essays how auditors can utilize process mining in their audits. Specifically, the focus of the dissertation and its contribution will be on applications of process mining to internal controls. The first essay develops a methodology that illustrates how process mining can be used to test internal controls to provide an overall risk assessment of the internal control system for a business process. Regulatory compliance requirements in the area of Internal Controls such as the Sarbanes Oxley Act force firms to report on the effectiveness of their internal controls. Auditors are required to assess the effectiveness of the firm's internal control system and issue an opinion. Traditionally, auditors use qualitative methods to complete this process. However, this is far from an objectively efficient method to measure controls consistently and effectively. Moreover, considering the consequences of the failure to accurately measure the effectiveness of internal controls and assess its risk, auditors should be eager to embrace a more formal internal control assessment process with quantitative outcomes. This conceptual framework was tested on a set of data that relates to the procurement process obtained from a national not-for-profit organization. The results

have found several internal controls to be lacking in different areas of the procurement process.

With the large number of transactions being executed on a daily basis, auditors are facing increasingly difficult challenges in detecting and investigating anomalies and exceptions. The second essay proposes a methodology that provides auditors with guidance on the use of process mining in conjunction with existing analytical procedures to identify exceptional transactions that would require further investigation. This solution allows auditors to focus on process instances that are likely to be considered high-risk, reduce the risk of failing to detect material misstatement, and enhance audit effectiveness. Furthermore, the identification and prioritization of such risky process instances help with the problems that result from population testing, such as information overload.

The third essay proposes a conceptual methodology that illustrates how a rule-based process mining technique can be used to provide continuous monitoring of controls for a business process. The periodic nature of auditing and monitoring creates a time-delay between the occurrence of important business events and the analysis of the events. Advances in technology provides opportunities to reduce the time delay between the occurrence and analysis of a business process event. By significantly reducing the time-delay, the created information becomes more valuable since it allows for additional management control and assurance activities. The conceptual framework is demonstrated using event logs from the procurement process obtained from a national not-for-profit organization. The continuous monitoring layer of the framework has the capability to detect and prevent multiple violations.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my mentor and dissertation chairman Dr. Miklos Vasarhelyi for inspiring me, encouraging me, and guiding me to become a better researcher. Without his help and support, this dissertation would have not been possible.

I would also like to gratefully acknowledge the contributions and guidance of my dissertation committee members, Dr. Helen Brown-Liburd for her insightful feedback and guidance, Dr. Hussein Issa for his continuous support, encouragement and mentorship throughout my PhD journey, and Dr. Rajendra Srivastava for his appreciative comments and constructive feedback.

I am indebted to my many colleagues and friends for assisting me throughout my time at Rutgers and making it an enjoyable and insightful journey. I would like to specially thank Abdullah Alawadhi for his inspiration and help, and Ahmad Alqassar for his companionship along this journey. I would also like to thank my dearest friends Tiffany Chiu, Andrea Rozario, Jamie Freiman, Yunsen Wang, Zhaokai Yan, Eid Alotaibi, and Zamil Alzamil. I will always cherish your friendship and academic support.

Lastly, I greatly acknowledge the love and encouragement of my family, especially my parents, I wouldn't have achieved this without them. I would like to especially thank and give my deepest appreciation to my wife Hibba Ihmeidan, who stood beside me. Thank you for your love, support, and continuous faith in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Setbacks in the business world such as the scandals in Enron, Tyco, Adelphia, Peregrine, and WorldCom, along with the recent financial crisis have pressured the accounting profession to introduce more rigorous auditing practices. As a response, new legislations like the Sarbanes Oxley Act of 2002 (SOX) and the Basel II Accord of 2004 were enacted. Under section 404 of SOX, publicly-traded companies are expected to establish internal controls for financial reporting (ICFR, and both management and their auditors are required to report on the effectiveness of ICFR. Auditing Standard No.12 (AS 12) indicates that auditors should assess the risk of material misstatement, which includes both inherent and control risks. The inherent risk refers to the susceptibility of an assertion to a misstatement before any control is exercised. Control risk expresses the risk that a misstatement will not be prevented or detected on a timely basis by the company's internal controls. Among the procedures proposed by AS 12 to assess the risk of material misstatements is obtaining an understanding of internal controls over financial reporting and performing analytical procedures. The process of obtaining an understanding of internal controls includes evaluating the design of controls that are relevant to the audit and determining whether the controls have been implemented. Thus, auditors generally conduct "walkthroughs" that include inquiry of appropriate personnel, observation of the company's operations, and inspection of relevant documentation.

In today's digital economy, technology allows for different tools to validate information about companies and their business processes. One very powerful recently

developed tool is process mining. Process mining is a methodology developed by computer scientists and was utilized in the industrial and software fields to help with discovering, monitoring, and improving actual processes by extracting knowledge from unstructured data sources. In recent years, the accounting literature suggests that the audit profession can make use of such a tool. Implementing process mining in the audit process allows auditors to test the entire population, rather than using the traditional sampling approach, and base their opinions on an objective data source in the form of meta-data from the company's ERP system.

1.1. HISTORY AND EVOLUTION OF PROCESS MINING IN AUDITING

In the early 90s, areas of Workflow Management and Business Process Reengineering (BPR) attracted research attention in the form of business process redesign and innovation (Hammer and Champy 1993, Davenport 1993, Datta 1998). Logically, models of existing processes should be attained before performing BPR. Although some BPR research assumes that models of organizational processes are known before reengineering, others do recognize the difficulty and cost of their extraction (Hammer and Champy 1993, Davenport 1993, Datta 1998).

Business processes specify the way resources of an enterprise are used (Agrawal et al. 1998). Each process is usually comprised of a set of activities that may or may not be dependent on each other. The main task of workflow systems is to ensure that all the activities are performed in the right order and the process terminates successfully (Leymann and Altenhuber 1994, Agrawal et al. 1998). Agrawal et al. (1998) introduced

the idea of applying process mining in a workflow management context (van der Aalst et al. 2003).

Reliable information about the operation of any organization is essential for stakeholders in order to make a variety of decisions. The objective of auditing is to provide stakeholders with reliable information by validating information generated from business processes. Traditionally, auditors select samples of the population to assess the operating effectiveness of process controls. With the development of technology in general and enterprise resource planning (ERPs) in particular, detailed information about processes in the form of event logs became increasingly abundant. With such developments, current research (van der Aalst et al. 2007, 2011, 2010; Jans et al. 2011, 2013, 2014) proposed and tested the use of process mining in the auditing domain.

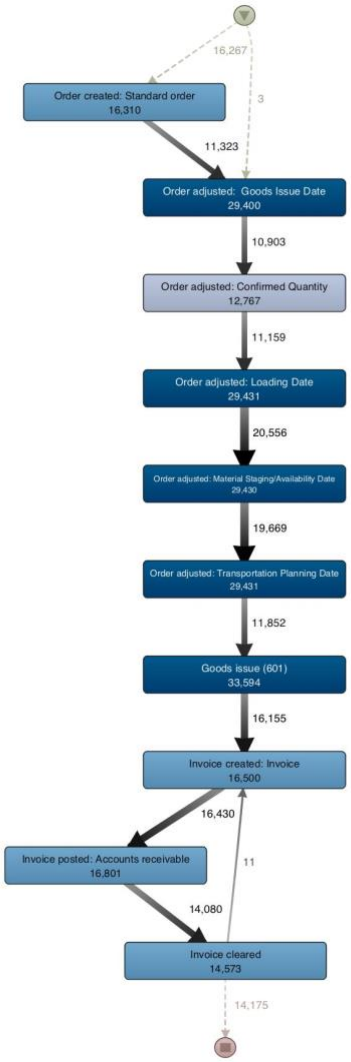
Van der Aalst et al. (2010) introduced an auditing framework by the name of Auditing 2.0. The framework indicates that two types of data can be extracted from the information system, current and historical. There are also two kinds of models presented in the framework, De Jure and De facto. De jure models are the required models, whereas de facto models describe what is happening in reality. The de facto models are reached by using process mining techniques to extract a Petri net that models behavior in the event log found in historical data. The auditor can then perform multiple tests to validate the company's process. The auditor can check if the historic data in the event log conforms to the desired model in order to detect deviations, locate and explain them, and measure their severity (Rozinat and van der Aalst 2008). The auditor can also compare de jure and de facto models in order to analyze the differences. Finally, auditors can diagnose de facto models by using model-based analysis techniques to check for deadlocks and other

anomalies. Van der Aalst et al. (2010) proposes a new and effective way of conducting audits but the main concern is blending the obligations of internal and external auditors. For example, things like extending de facto models interferes with the independency of the external auditor.

1.2. APPLICATION OF PROCESS MINING IN INTERNAL CONTROLS

The difficulty in understanding and evaluating the internal control environment in today's business processes can be addressed by the utilization of process mining techniques. Process mining can address the problem that most internal control experts face, which is having minimal information about what is actually happening in the business processes (Caron and Vantheinen 2012). Management is responsible for determining the standard process flow for any business process that is part of the organization's operations, and employees need to abide by those ideal process flows. Figure 1 shows an example of the standard process flow for an order-to-cash process.

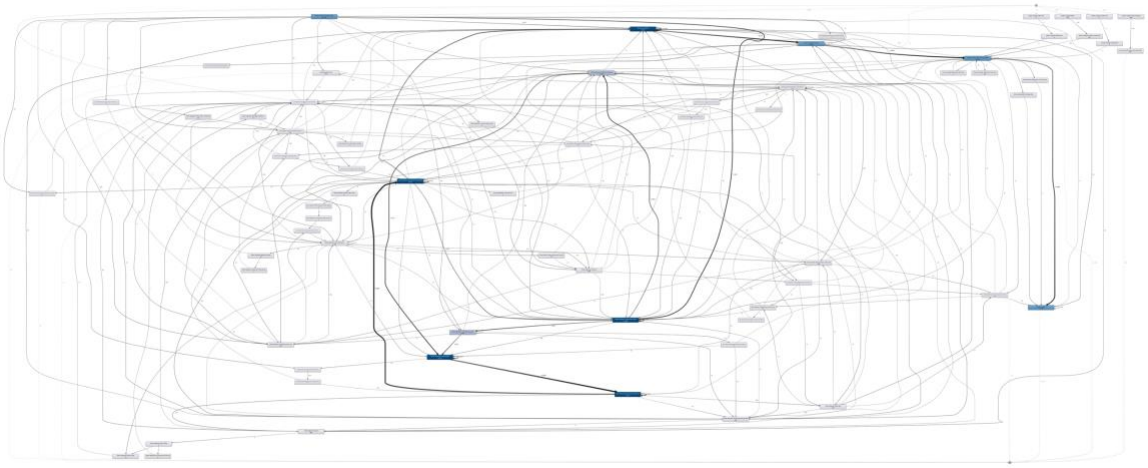
Figure 1. Standard Process Flow for order-to-cash process



In reality however, employees and managers override business rules and deviate from the ideal process design out of operations necessity. For example, a bank manager must have leeway to override certain rules such as credit limits for a customer that might be considered a strategic client or might have other accounts with substantial balances. Figure 2 shows the actual process flow of the same order-to-cash process found in Figure 1. This illustrates how process mining is a powerful tool not only in assisting auditors in

understanding how processes actually operate, but also allow them to focus on specific activities which pose the greatest control risk (Jans et al. 2013).

Figure 2. Actual Process Flow for order-to-cash process



In addition to understanding a business process, Caron and Vantheinen (2012) discuss the advantages that process mining has for a more efficient and effective control environment. These advantages include:

1. *Gaining detailed and objective information on the business process*

   Due to the flexibility that is demanded by business necessity and accommodating customers' needs, deviations from the standard process will occur. Therefore, process discovery can help auditors in understanding the reality of the business process, and highlight any weaknesses or concerns in internal controls.

2. *Obtaining high levels of assurance*

   Process mining analyzes the full population of data instead of resolving to traditional audit procedures utilizing sampling techniques, thereby, offering high levels of assurance. Moreover, since all instances are examined, the risk of the

evidence failing to uncover misstatements can be significantly reduced.

3. *Gathering strong evidence*

   The strength of evidence gathered using process mining is expected to be strong. The main reason is that process mining allows auditors to rely on data that is independently produced from the system and the ability to examine the whole population instead of a sample.

This dissertation will demonstrate how auditors can utilize process mining in their audit engagements. Specifically, the focus of the dissertation and its contribution will be on applications of process mining in internal controls. The first essay examines how process mining can be used to assess the risk and evaluate the effectiveness of controls over financial reporting. Both the design and operational effectiveness of controls is examined to assess the risk. The second essay attempts to resolve the issue of the large number of false positives associated with full population process mining analysis. It proposes a methodology by which process mining can be used to prioritize and rank suspicious process instances. The third essay demonstrates a continuous process monitoring methodology that reduces the time delay between the occurrence and analysis of a business-related event. The methodology is based on implementing an abstraction layer on top of the business process, and a rule-based process mining technique.

CHAPTER 2: THE APPLICATION OF PROCESS MINING IN INTERNAL CONTROL RISK ASSESSMENT

2.1. INTRODUCTION

Due to the setbacks in the business world, the Sarbanes-Oxley Act (SOX) was introduced to ensure that publicly-traded companies have adequate internal controls. Under SOX section 404 companies were expected to establish internal controls for financial reporting and assess them via auditors to ensure their effectiveness. However, auditors routinely fail to detect material weaknesses before a restatement. One reason is that auditors inaccurately assess control risk assessment by misclassifying the severity of identified internal control deficiencies due to complexity in judging the materiality and likelihood of potential related errors (Aobdia et al. 2016).

When it comes to assessing the internal control system of an organization, most guidance provided recommends a top-down approach to internal control evaluation. That means that auditors should begin with an identification and assessment of risks at the financial statement level, and then move down to the significant account and disclosure level to determine whether controls have been placed in operation to address those risks. This includes identifying relevant business processes affecting the significant accounts, and control objectives specific to the organization that must hold for each process, as well as continuously assessing the risks, and the design and implementation of controls in order to prevent or detect the occurrence of the identified risks.

By taking this approach, auditors can more easily identify items at the financial statement level that are highly risky and have significant or material balances, which must

be addressed from a control perspective to ensure that financial statements are fairly presented. Likewise, when risks are minimal and account balances are relatively insignificant, management would not have to expend extensive resources over controls for those reporting areas. The framework that is recognized by regulatory bodies and auditors as a de facto standard for assessing internal controls systems is the one offered by the Committee of Sponsoring Organizations of the Treadway Commission (COSO). However, the COSO framework provides high-level guidance for internal control and does not provide the detailed control objectives and measurements required in the design of audit tests. Also, it does not address the specific risks and complexities of information technology (IT) (Chang et al. 2014).

Organizations today have implemented some type of computerized information systems to increase their efficiency and to cope with the changes in the business environment. Internal controls adopted under non-computerized environments have become less relevant and effective in preventing or detecting errors on a timely basis for an organization implementing advanced IT accounting system. Therefore, management has developed and adopted new and valid internal control tools and procedures. Consequently, auditors need to adopt a valid and appropriate framework to assess the effectiveness of the internal control system under an advanced IT environment (Hwang et al. 2004).

Traditionally, the way auditors evaluate controls is by using conventional qualitative methods. These methods could be a mix of checklists, questionnaires, flowcharts, and tests of transactions. However, research has shown that such methods are considered to be insufficient and the assessments generated by qualitative methods alone are insufficient for developing comprehensive internal control evaluation models (Yu &

Neter 1973; Cushing 1974, 1975; Mock & Turner 1981; Bierstaker and Wright 2004). Moreover, auditors not only have to evaluate controls on the basis of whether controls are implemented or not but also have to measure and assess the operational effectiveness of those controls. Therefore, it has become necessary for auditors to apply different techniques and tools to test controls, especially under an advanced IT environment. This is because most transactions are processed using programmed procedures and their related data and evidence are stored electronically. Hence, audit tools and techniques that are applied to a traditional accounting system are inapplicable and inefficient (Hwang et al. 2004). Current technology allows auditors to take advantage of advanced audit tools and techniques that enable them to examine the entire population, leading to a more effective and efficient audit.

COSO defines the Internal Controls as a "process" designed to provide reasonable assurance regarding the achievement of objectives in effectiveness and efficiency of operations, reliability of financial reporting and compliance with applicable laws and regulations. Process mining is a tool that allows auditors to model the actual workflow of a business process from activity logs stored in the organization's information system. The main idea behind process mining is "to diagnose processes by mining event logs for knowledge" (van der Aalst and de Medeiros 2005). With the vast number of transactions and complexity of modern information systems, it is challenging to assess the design and compliance of controls in real world scenarios. Therefore, the motivation behind this study is that there have been many studies in the past that focused on building frameworks for internal controls (Bailey et al. 1985), automating internal controls (Alles et al. 2006), and evaluating internal control systems using belief function (Mock et al. 2009). However, few

studies provide a methodology for assessing internal control risk in a formalized and systematic way.

This study introduces a framework that assists auditors in quantitively evaluating internal controls risk assessment. Auditing Standard No. 5 requires auditors to gather evidence to support their opinion about the internal controls over financial reporting for an organization (PCAOB 2007). The auditor must obtain evidence about the effectiveness of selected controls concerning all relevant assertions. This requires that the auditor test the design and operating effectiveness of controls. Aslo, when certain conditions are met, auditing standards (AS No. 5) permits auditors to use a benchmarking strategy for fully automated application controls. Process mining techniques provide strong evidence on the effectiveness of controls (Jans et al. 2011; Jans et al. 2013; and Jans et al. 2014). Therefore, the aim and contribution of this study are to propose a methodology to measure and asses internal control risk. Specifically, the study will present a conceptual model that illustrates how process mining can be used to test internal controls, and then be used to provide guidance for assessing control risk. The system attempts to run tests on a dataset relative to a specific audit function, produce results, and based on those results, provide a formalized measure for the effectiveness of the internal control system.

There has been a call in the auditing literature for developing a baseline of control effectiveness measurement. Regulations require auditors to assess internal control risk both in terms of implementation and operation. In a traditional audit, auditors rely on the use of sampling due to the labor and time intensiveness of manual testing. In contrast, advanced audit tools, such as process mining, would consider the whole population of transactions in testing. The consideration of the whole population of transactions in testing can enhance

the effectiveness of an audit and increases the probability that material errors, omissions, fraud, and internal control violations may be detected (Chan and Vasarhelyi 2011). Additionally, when auditors run tests and analytics on a given data set relative to a specific audit function, results or exceptions generated from these tests are usually investigated to see why specific internal controls were not in place or were not effective. However, these same results can be used not only for investigative purposes but also to provide a measure for the effectiveness of an internal control system from an operating perspective.

Current literature lacks studies that address the issue of objectively measuring the level of adequacy of internal controls. This could be the result of the scarcity of feasible real-world data (Amat, 2002). This study is motivated by the difficulties facing auditors in identifying weakness and deficiencies in the internal control systems due to the added complexity of business processes in today's digital economy. Therefore, this study contributes to the auditing literature by demonstrating how process mining can be used as a tool to identify deficiencies in the internal control system and proposes a framework that auditors can use to quantify and objectively assess controls risk. In this study, the conceptual model was tested using a set of data that relates to the procurement process obtained from a national not-for-profit organization. Results demonstrate a lack of controls in several areas of the procurement process.

This paper is organized as follows: section (II) will provide a background and literature review on the importance and assessment of internal controls. This will be followed by section (III), which describes the methodology and general framework developed in this study. Section (IV) will demonstrate the methodology on a specific business process, and describe the data used, the analysis, results, and discussion. Finally,

section (V) presents the conclusion.

2.2. LITERATURE REVIEW

2.2.1. INTERNAL CONTROLS AND PROCESS MINING

Companies and auditors have long considered internal controls to be of vital importance (e.g., Mautz and Sharaf 1961; AICPA 1983; COSO 1992). For instance, the Security and Exchange Committee (SEC) has requested as far back as 1941, most auditors to consider a company's internal controls in planning an audit (SEC 1941). Auditors and researchers have tried for over 40 years to develop approaches for assessing risks and evaluating internal control systems. They strived for methodologies that would be rigorous, systematic, and tractable in practice. One of the earliest of these studies was Bailey et al.'s (1985) The Internal Control Model (TICOM) for designing, analyzing, and evaluating internal control systems. TICOM required its users to code agents and tasks, and use a query processor to analyze the model. "A bilogic-directed graph showing both control and data flows" was used for internal representation (Bailey et al. 1985). The main shortcoming of TICOM was that its use was not practical, and very burdensome to represent the system in the programming language (PASCAL), and auditors were not familiar with the logic embedded in graph representations (Borthick 2012). Ever since Bailey at al.'s (1985) study, other researchers have developed mathematically based frameworks, but they were generally not adopted because they were not applicable in practice.

An even earlier study by Cushing (1974) introduced a mathematical technique of a simple stochastic model based on reliability theory adapted from the field of reliability engineering to evaluate the design and effectiveness of an internal control system. The

model provides a means of computing the reliability of a process, which is the probability that the process will be completed with no errors. Cushing (1974) describes a means of representing internal control in mathematical terms, and demonstrates how such mathematical representations may be useful to controllers and auditors in designing and evaluating internal control systems

Any firm's internal control system is affected by the different decisions that the board of directors or management have to make when designing and implementing it. AS No. 5 describes internal control evaluation as a risk-assessment process, and requires both the firm and its auditor to assess various aspects of the firm's internal control system to provide reasonable assurance regarding the effectiveness and efficiency of operations, the reliability of financial reporting, and compliance of the organization with laws and regulations (COSO, 2013).

Auditing literature has long recognized the importance of assessing control risk and evaluating the effectiveness of internal controls because of what it provides to any organization (Mautz and Sharaf 1961). Prior research has called for research into quantitative methods for evaluating the effectiveness of internal controls. Both a stochastic model (Yu & Neter 1973), and a reliability model (Cushing 1974) were developed and improved upon by several researchers (e.g., Grimlund 1982; Srivastava and Ward 1983; Srivastava 1986). Unfortunately, research on internal control assessment methods has been somewhat scarce in the past decade (Mock et al. 2009).

Although, in today's general move towards a digital economy within the business world, organizations are heavily relying on network integrated information systems, such

as ERP systems. This increased reliance on technology has increased the complexity of business processes due to the high number of transactions and simultaneous processes (Kogan et al. 2010). This complexity is naturally extended to the auditors and can be overwhelming without sound guidance (Tuttle and Vandervelde 2007). To deal with this complexity in assessing the effectiveness of the internal control system, the PCAOB requires companies to adopt an internal control framework by which its practices can be assessed, and mentions the COSO as one. Most companies adopt COSO, but other frameworks can also be used.

Auditors have utilized several techniques in assessing the effectiveness of internal control systems over the last three decades. These techniques help auditors in collecting and organizing raw data for evaluating internal controls (Cooley and Hicks 1983). Related studies have found these techniques to be effective in delivering relevant task information to junior auditors, and further improve their task performance (Graham, 1993). Also, another study found that flowcharts can assist auditors in constructing a precise and comprehensive mental model of complicated systems (Brewster 2008). However, such traditional qualitative judgment methods used by management and auditors, such as "High," "Moderate," and "Low" rather than quantitative judgment are insufficient when developing comprehensive internal control evaluation models (Yu & Neter 1973; Cushing 1974, 1975; Mock & Turner, 1981; Bierstaker and Wright, 2004; Mock et al. 2009; Norman et al., 2009).

As a solution for today's business process complexities, process mining has been proposed by researches as an aid to help in these challenges. Process mining is a tool that analyzes event logs extracted from the organization's information system. Although

process mining has been around for a while, it is relatively new to the accounting literature. With such developments in IT and process mining methodologies, current research propose and test the use of process mining in the auditing domain (Van der Aalst et al. 2010, 2011; Jans et al. 2011, 2013, 2014). Jans et al. (2013) also argues the case for adopting process mining in auditing:

1. It enables the auditor to examine the whole population of transactions, rather than the current sampling method
2. The transaction entries are generated automatically from the ERP system, thus eliminating the dependency on potentially subjective data provided by the auditee

Basically, process mining can be used to model the design of business processes and find evidence that controls are operating effectively (Agrawal et al. 2006). Also, process mining can be a good quantitative representative on the level of risk an organization has based on the frequency of violations and the material impact of violations on the financial statements (Caron et al. 2013).

2.2.2. PROCESS CONFORMANCE AND RISK ASSESSMENT

One of the key determinants of the auditor's ability to appropriately plan and conduct the audit is affected by his or her ability to effectively analyze operations in the form of business processes (Carnaghan, 2005). Hence, current auditing standards emphasize the importance of auditors' understanding of the operation of an organization by performing a risk assessment. Since the majority of organizations implement an information system for their operations, such as SAP, process mining can be increasingly

used (Wu et al. 2007; van der Aalst and De Madeiros 2005). Hakvoort and Sluiter (2008) have determined that process mining should not only be used in the planning phase, but should also be used in executing the audit.

In order to assist auditors in the audit risk assessment process, a suitable approach is a process conformance checking technique (Hakvoot and Sluiter 2008). Process conformance checking means that every process instance or transaction is checked against a prescribed process model. If the process instance does not match the prescribed process model, then this deviation could be indicative of a control failure and an undesired exception.

In audit practice and theory, one of the most and widely accepted concepts is the ability of the client's internal control system to generate reliable financial information and safeguard assets. Therefore, auditors are required to assess control risk, which is the process of identifying internal controls and evaluating their effectiveness. The primary purpose of designing a system of internal controls is to provide reasonable assurance regarding the achievement of management' s objectives as it pertains to the reliability of financial reporting, the effectiveness and efficiency of operations, and compliance with applicable laws and regulations (Arens et al. 2003).

In the planning phase, auditors assess whether controls are in place by examining the design of the internal control system and determine whether they can rely on them. If controls are in place and the design of the internal control system is adequate, then auditors must test the effectiveness of the internal controls to justify the reduced control risk and the amount of audit evidence to be accumulated. Test of controls can be accomplished by

using the process conformance checking technique since event logs can be analyzed and checked to determine whether all required steps have occurred in the correct order (Hakvoot and Sluiter 2008).

2.2.3. APPLICATION OF PROCESS MINING THROUGHOUT THE AUDIT CYCLE

Auditors are likely to benefit from process mining throughout the audit cycle (van der Aalst et al. 2010). By using process mining, auditors gain a clear understanding of the client's business processes and its environment, traditionally accomplished via walkthroughs. Traditional walkthroughs may not provide auditors with a complete picture of the entity's business processes. Also, exceptions may not be captured in traditional walkthroughs since only very typical ways of performing a process are discussed. Moreover, using process mining allows auditors to identify and assess business risks and test for internal control weaknesses. Process mining can be applied by auditors throughout the audit cycle, including the planning, fieldwork and reporting stages.

In the planning stage, auditors start by gaining a general understanding of the overall processes and conduct risk assessments to identify any potential material weaknesses. Process mining can be used in this stage by performing process discovery, which helps auditors further understand the business processes, and identify any potential risks in order to create an effective audit plan. In the audit fieldwork stage, auditors may utilize process conformance checking techniques to perform tests of controls and reduce the planned substantive testing for related accounts (Hakvoot and Sluiter 2008). Auditors

can also use process mining as a supplement to analytical procedures and other forms of evidence collection. Since process mining analyzes the entire population of data, the strength of evidence gathered is considered high. Finally, in the reporting stage, auditors may rely on the visualization of process models and discoveries to present results to management.

2.3. METHODOLOGY & FRAMEWORK

When assessing control risk and the overall effectiveness of internal controls over financial reporting, there are different levels that auditors take into consideration. These levels include significant accounts level and business process level. Information gathered from evaluating individual controls is valuable in facilitating the process of identifying any significant weaknesses existing in the internal control system, and for optimizing the value of internal control investment (Mock et al. 2009). The generic measurement of internal control effectiveness model developed in this study is part of a model of risk assessment that auditors would use as implemented under Auditing Standard No. 5 (PCAOB 2007). The generic risk assessment model consists of a financial reporting part and a business process part. However, for the purpose of this study, the paper will only focus on the business process part.

The business process part consists of the management assertions concerning internal control over the financial reporting system related to the significant accounts, risks associated with these assertions, and the control procedures implemented to mitigate these risks. Thus, internal controls are designed to control risks specific to management's

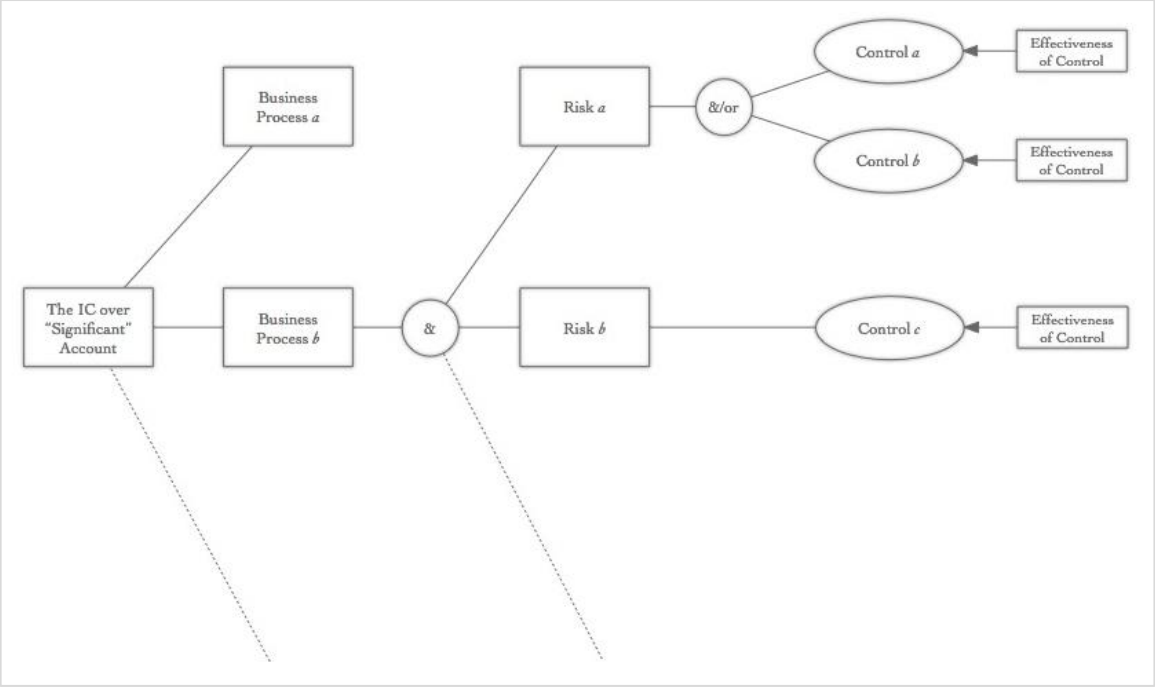
assertions concerning the accounting information system effectiveness. Broadly speaking, for each management assertion, there are several potential risks, and for each risk, there may be more than one internal control to mitigate the risk. One or more risks may threaten each assertion. Thus, for a system to be effective, risks of not achieving assertions need to be mitigated by one or more controls.

Since the passage of SOX, the importance of using frameworks to guide the assessment of internal controls has dramatically increased. The use of a framework results in more comprehensive, reliable, and complete assessments. However, achieving these goals in today's IT intensive environment is difficult without a control framework that conceptualize the important aspects of internal control within an IT context in a complete and logically consistent manner. The COSO framework, which is recognized as a de facto standard by regulatory bodies for realizing controls for financial reporting, focuses on high-level guidance for internal controls and does not provide the detailed control objectives that auditors need in the design and assessment of control testing. Moreover, the framework does not address the specific risks and complexities of IT. Without a comprehensive and conceptually sound framework, auditor can get overwhelmed by the complexity of modern systems. This suggests that the quality of assessing the internal control system depends on the conceptual model upon which a framework rests (Chang et al. 2014).

Process mining techniques are used to gather direct evidence on the operational effectiveness of a business process, and highlight any deficiencies in the internal control system, both from a design and operation perspectives. These deficiencies are scored based on the severity of the violation and significance of the controls in place to mitigate the risk. Figure 3 illustrates the general business process model and the assessment of the internal

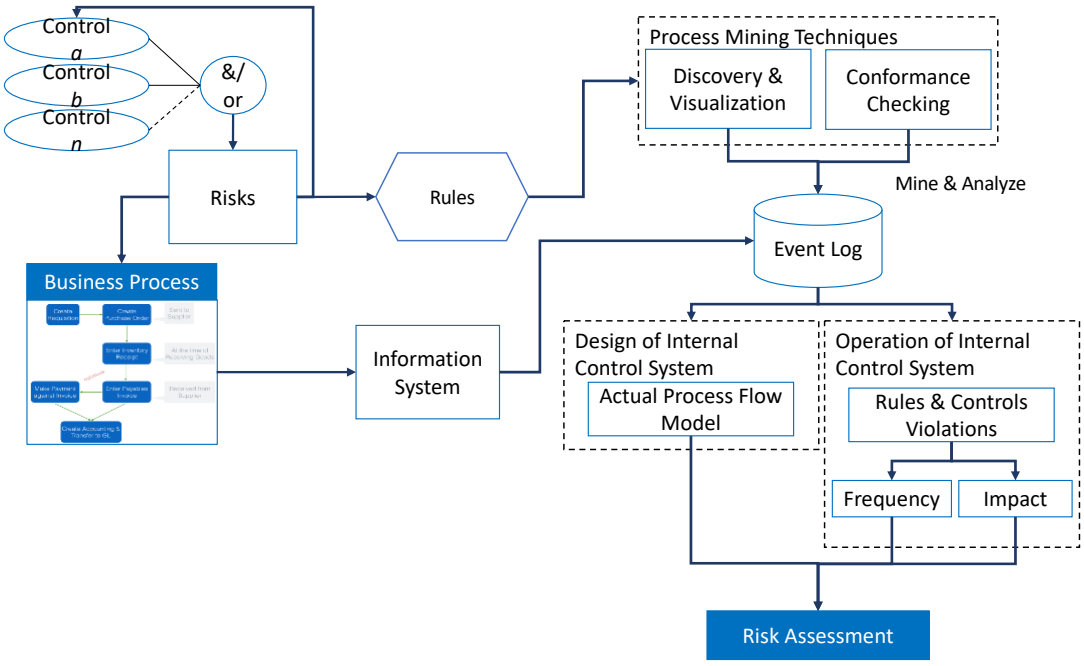
control system effectiveness.

Figure 3. Business process model and the assessment of the internal control system effectiveness



Any business process consists of many different controls implemented to mitigate the risk of fraudulent activities and ensure that the financial statements are represented fairly. There are two main aspects that auditors focus on when evaluating an internal controls system. The first aspect is evaluating the design and structure of the internal control system. The second aspect is evaluating the operational effectiveness of the internal control system. These two parts are shown in figure 4 where the internal control risk assessment framework builds on the general process model.

Figure 4. Internal Control Risk Assessment Framework



The framework consists of two main parts: First, process mining is applied to understand the business process and discover risks that were not highlighted by the traditional and standard What Could Go Wrong (WCGW) from the auditor's associate with it. This part focuses on the design aspect of the internal control system and investigating weaknesses from a structural level. Second, gather direct evidence on the effectiveness of controls by applying a rule-based conformance checking process mining technique that would serve as the bases for a quantitative risk assessment of ICoFR in an efficient and effective way.

2.3.1. INTERNAL CONTROL SYSTEM DESIGN

Traditionally, auditors assess the design through the inspection of existing process

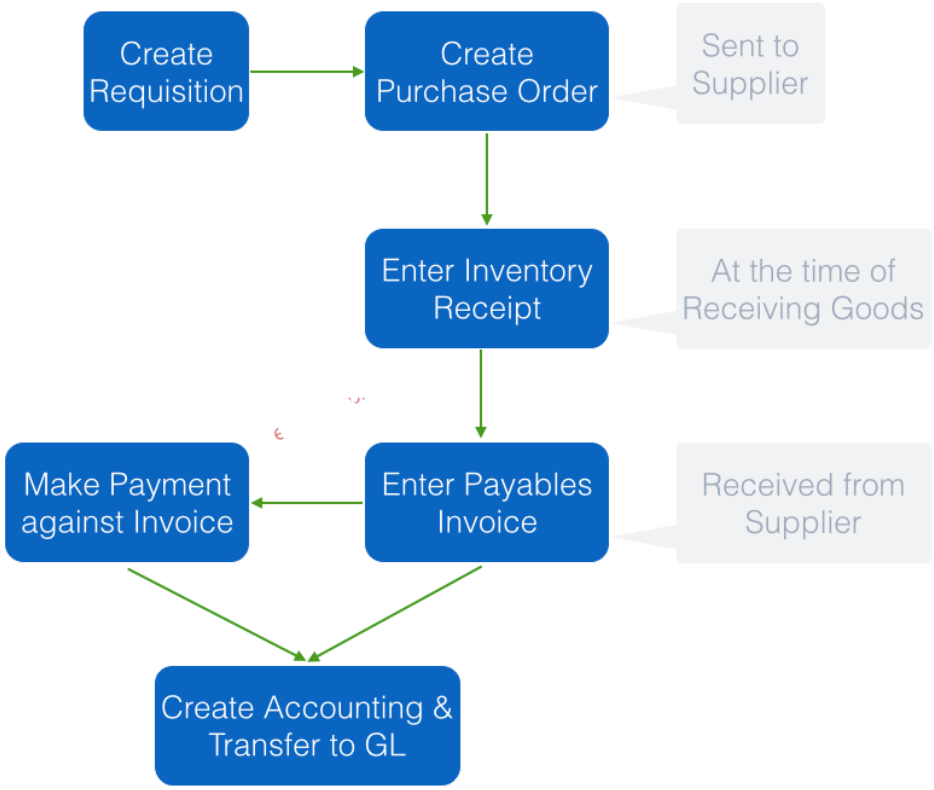
documentation and interviews with management and employees. After that, auditors identify different controls that need to be in place to mitigate risks associated with management assertions. The effectiveness of the internal control system is affected by how well it is designed to mitigate those risks. Evaluating the design of the internal control system can be facilitated using process mining. This study will demonstrate the utility of this solution and break it down into steps that auditors can follow in order to evaluate the design of the internal control system.

Process Flow Modeling

Part of understanding the business process and the actual way transactions are being conducted, the first component of the framework allows auditors to evaluate the design the internal control system for that specific process. This is achieved by modeling the actual process flow following a bottom-up approach. Using the organization's information systems, auditors need to initially gather the required data to model the process flow of transactions in the desired business process. This is done by constructing the event log of all recorded transactions found in tables stored in the information system.

Event logs of past activities are used to provide a baseline model on the actual process flow of the intended business process and highlight any deficiencies or weaknesses in the design of the internal control system. Auditors are able to visualize the design of the of the internal control system and the steps that are taken throughout the process to complete each transaction. Figure 5 shows the most common model, and in this case, the ideal steps that should be followed and in the same order to complete a routine transaction. Any transaction that follows this path is considered acceptable from a process perspective.

Figure 5. Standard Process Model for a Routine Purchasing Transaction



Process Discovery

Process discovery allows auditors to expand that view and visualize every path that was taken by the transactions of the business process. Figure 6 provides a detailed process model that includes all the paths in the event log. This component of process mining allows auditors to analyze the event log in order to discover how the actual process is carried out and whether it differs from the designed process model. This is done by analyzing the number of variants, which are the unique sequence of activities in the event log. Since

transactions of business processes in an organization is conducted in many different ways depending on the circumstances, the analysis of the event log will generally produce many variants. Therefore, it is imperative that the auditors distinguish between the ones that are carried out of necessity rather than violations of internal controls.

In addition, visualization is a tool that is used in the process discovery component and provides auditors with a way to discover deviations and assess the design of the internal control system. Auditors can instantly notice some notable variants where controls are violated. For example, auditors can discover from the model in figure 4 that 6 cases started with an Invoice Receipt (IR) activity instead of a Create Purchase Order (Create PO) activity as required by the business rules and the standard process. Also, 91 Create PO activities were directly followed by Goods Receipt (GR) activity without being authorized. Figures 7 and 8 show these violations in detail respectively. These deviations of the ideal way of completing transactions can be discovered using process mining and ultimately assist auditors in assessing the design of the internal control system by detecting the occurrence of such deviations.

Figure 6. Detailed Process Model for All Purchasing Transaction

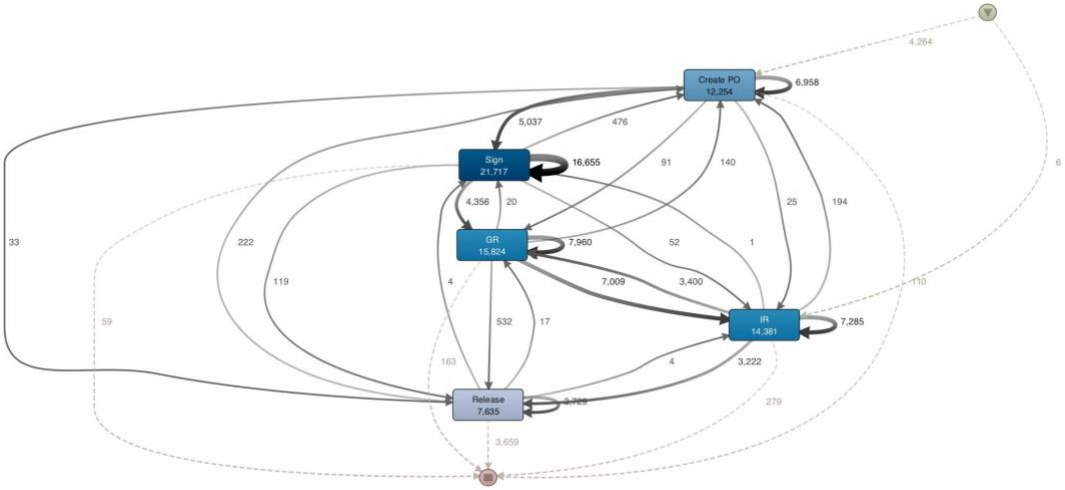


Figure 7. Violation Example from Process Discovery: Incorrect Start of Transaction

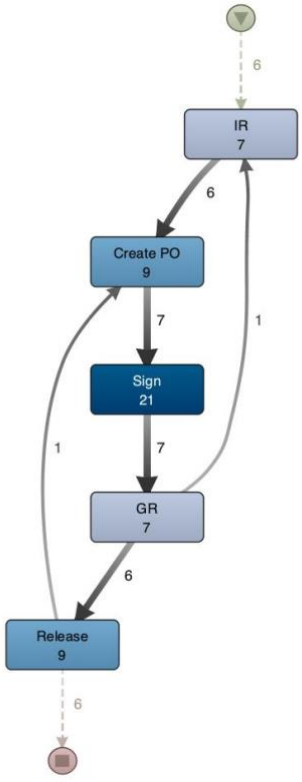
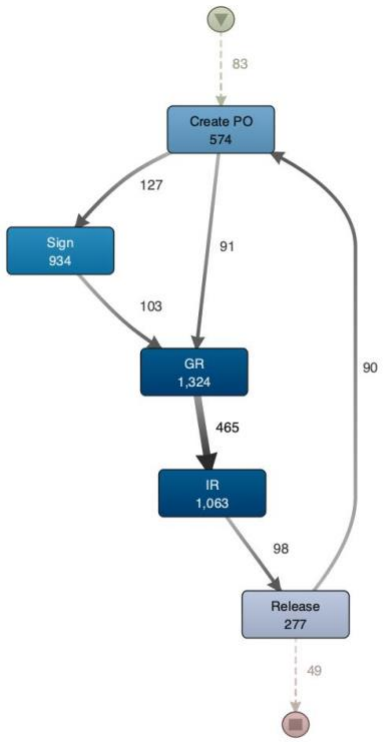


Figure 8. Violation Example from Process Discovery: Incorrect Order of Transaction



Process Deviations

As part of process discovery, auditors contrast process instances or transactions with the ideal designed business process model. If a process instance does not follow the exact path of the ideal process model, then it is considered a process deviation. The deviations could be the result of missing a key control or activity, a redundant activity, or an activity not in the right order (Chiu and Jans 2018). From analyzing these process deviations, auditors can assess the internal control system from both a structural level (investigating whether controls are implemented or not) and an operational effectiveness level (how frequent are controls violated). Process mining, like any other analytical

procedure, when applied to the whole population results in a large number of deviations and could overwhelm the auditor with a flood of false positives. Therefore, auditors need to validate whether the audit-relevant information generated by process mining is really indicative of weaknesses in the internal control system.

Risk and Controls

When evaluating the design of the internal control system, auditors need to identify the different controls for the business process that need to be implemented by the organization to attest management assertions and mitigate risks. For example, authorization of POs is a control designed by management to prevent potential misstatements, both intentional and unintentional. Generally, there are multiple levels of risks that auditors consider. Auditors can assess the level of risk based on the industry that the organization is in. Auditors can also specify certain risks associated with the business process being audited. With all of these different risks, auditors come up with a list of WCGW that are relevant to the business cycle, and the appropriate controls that need to be implemented to mitigate those risks and attest management assertions. An example of a WCGW is that an employee may have the ability to initiate, authorize and record a transaction or may have custody of assets within the process, such that they are able both to perpetrate and conceal an error or irregularity. Therefore, a segregation of duty control needs to be implemented and working effectively to ensure that the activities that management want segregated are actually segregated in all instances throughout the process. Consequently, a control should be in place to prevent or detect when the same person creates the PO and performs, for instance, the activities of GR and IR. Management is responsible to ensure that employees do not perform any inappropriate combination of activities. Process mining is a very

powerful tool in detecting breakdowns in the internal control system and highlighting deficiencies and areas of risk whether they are from a design perspective, or an operational effectiveness perspective.

The identification of risks and controls can be facilitated by process flow modeling and process discovery analysis. The event log of the business process can be examined to highlight areas of risk and deficiencies in controls, or where controls are absent. Therefore, auditors can use different process mining techniques to gather knowledge about the business process and to evaluate the design of the internal control system from a structural level. The controls selected in this study are based on the business rules as explained by the data provider, auditing literature, and industry standards.

2.3.2. INTERNAL CONTROL SYSTEM OPERATIONAL EFFECTIVENESS

The second part of evaluating the effectiveness of the internal control system is evaluating its operational effectiveness. Auditors need to have reasonable assurance that the internal control system is operating effectively over the audited period. This is achieved through tests of controls. General and application controls that are used to meet the objectives of a business process can be tested through applying process mining techniques.

The methodology presented in this study to achieve the second part of evaluating the internal control system of a business process could be used as a baseline for control effectiveness measurement to assist auditors in evaluating the risk of the internal control system of an organization in a formalized and effective manner.

Once management assertions' risks are identified and their controls are

implemented, the next step is assuring that controls are operating effectively. The step involves applying rule-based conformance checking process mining techniques as a methodology for acquiring direct evidence on the controls' compliance. The process mining technique applied in this study is based on a rule-based system that is comprised of a set of IF-THEN rules that classify violating transactions as exceptions (Abraham, 2005). In this step, the whole population of data, such as transactions from the current fiscal period, rather than a sample, is tested against the rule-based system in order to identify exceptions. The rule based-system allows the user to run tests on a data set relative to a specific business process and produce results that reflect the effectiveness of the internal control, not just if the control is implemented by the firm or not.

The purpose of implementing a rule-based system is the simplicity it offers and its understandability by human users. Rule-based systems can be easily modified and are easily flexible to adapt to any changes. Furthermore, because of their simple logic, they are relatively easy to implement in practice.

Assess Effectiveness of Controls

This step highlights the deficiencies and exceptions discovered using process mining. process mining is used to analyze the event log in order to discover how the business process is actually carried out. This allows auditors to match and compare the discovered processes with a benchmark, enabling the identification of deviations that have taken place due to the necessities of operations or the violation of controls. Once these exceptions are generated, the effectiveness of control (EoC) can be calculated. EoC is equal to the number of violation indicators divided by the population. The equation below provides the general method for calculation EoC:

$$Effectiveness\ of\ Control = \frac{[n - x]}{n},$$

$x\ refers\ to\ the\ number\ of\ exceptions\ generated,$

$n\ is\ the\ total\ nmber\ of\ objects\ tested$

The EoC is an indicator of the plausibility that deficiencies in the control result in more than a remote likelihood that a material misstatement or fraudulent activities will not be prevented or detected. According to the PCAOB, there is a hierarchy of possible deficiencies: control deficiency, significant deficiency, and material weakness. There are two possible states of an internal control system; either it is effective or ineffective. Moreover, if the internal control system is ineffective then there are three possible conditions of ineffectiveness: deficiency, significant deficiency, and material weakness (Mock et al. 2009). These conditions depend on how severe the deficiency is and the level of tolerance that the organization has for materiality. The following set of EoC can be used to define the four levels of effectiveness and ineffectiveness of an internal control:

Effective Internal Control: EoC ≥ 0.95

Deficient Internal Control: $0.95 >$ EoC ≥ 0.90

Significantly Deficient Internal Control: $0.90 >$ EoC ≥ 0.80

Materially Weak Internal Control: $0.80 >$ EoC

Note that there is flexibility in these definitions and the stated ranges can be altered depending on the organization and the auditor's risk appetite, which can be defined as the level of risk and uncertainty auditors are prepared to accept. Additionally, despite using a simplistic probabilistic equation, the aim is to provide a basis for future calculations. The hope is that in future work, one can utilize different weighting schemas and various variables in enhancing the calculation methods and providing a more accurate measure.

Risk Assessment

The final step in the framework is assessing the risk of the internal control system for a business process. This is done by considering both the frequency and the impact of the violated controls on the financial statement. Auditors represent the results graphically on a risk map to determine which controls are riskier and ultimately have a more objective method in assessing control risk.

2.4. PURCHASING PROCESS APPLICATION

In this section, the general framework will be applied to the procure-to-pay cycle of a national non-profit organization. This procure-to-pay cycle will illustrate how auditors can use the general framework proposed in this study to evaluate the effectiveness of controls and assess risk. The results of each step in the framework will be detailed and analyzed.

2.4.1. DATA

A purchasing process from a national non-profit professional organization was selected for this study to assess process risk and evaluate the effectiveness of the internal controls. The process is based on a standard purchasing business process that is similar to many other businesses, which increases the generalizability.

The event log of the purchasing process analyzed in this study was extracted from the ERP system of the organization and the whole population of events was tested and analyzed. The event log dataset consisted of 4,270 purchase orders (POs) that were created between October 2014 and December 2016. Furthermore, one of the unique elements of this event log is that in addition to it containing information needed to reconstruct the paths of transactions of the purchasing process, it also has financial values related to each case or transaction. Regular event logs include information about the activity that was executed in each step of the process, the identity of the person who performed the activity, the time of execution, and other contextual information such as the role of the person involved in the activity. However, in addition to this information usually found in event logs, this specific purchasing event log includes the value of the PO, Goods Receipt (GR) value and the Invoice Receipt (IR)/payment value. This allows for additional analyses that traditional process mining cannot independently achieve. For example, auditors can assess internal control effectiveness by considering the frequency of a certain violation happening and its financial impact on the organization. This way, auditors can measure the severity of such violations and ultimately be able to take into consideration the principle of materiality for their risk assessment.

One of the challenges for auditors to implement process mining is the “materiality

principle" that guides current auditing practices (van der Aalst et al. 2010). By following this principle, auditors need to consider only a small sample subset of the data. And if they don't find any anomies or deviations, then they need not to take any further actions. However, since process mining typically allows auditors to examine the whole population, auditors will inevitably find more exceptions or violations that requires following-up. This will not only increase the quality of the audit, but also its time and cost. Therefore, the event log dataset used in this study and the solution proposed will help in minimizing the audit risk and minimize the effect of the materiality principle challenge by providing guidance to the auditor as to where to focus their efforts in further investigating violation that meet the minimum risk both in terms of frequency and severity.

2.4.2. ASSESSING THE DESIGN OF THE INTERNAL CONTROL SYSTEM

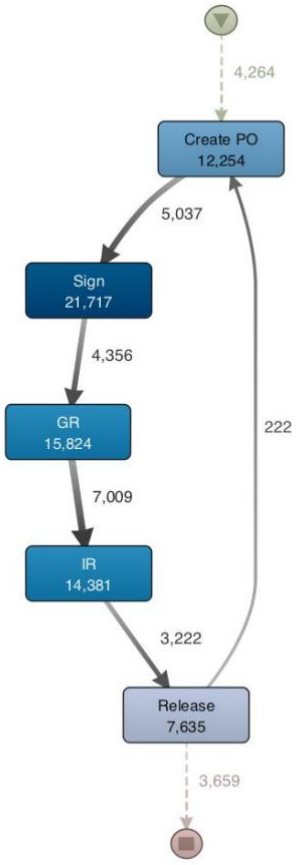
The first step in the framework is to model the process flow of the business cycle. In this case, it is the procurement cycle. To be able to model the process flow of transactions in the procurement process, an event log of all transactions is needed. The event log is extracted from tables in the information system of the organization and compiled together.

Once the event log is created, a preliminary analysis was conducted to identify the way purchasing activities were actually carried out in the organization. The commercial software application Disco is used for the process mining analysis (Fluxicon 2016). The default settings of Disco are based on the Fuzzy Miner algorithm of Gunther and van der Aalst (2007), and are applied in this study. This application basically filters typical issues encountered with large real-life datasets and then simplifies and visualizes complex

processes. Disco is an appropriate tool to model process flows and discover variants and deviations (Jans et al. 2014).

Analyzing the event log provides a preliminary understanding of the business process on hand. The descriptive statistics shows that there were 4,270 cases which represent unique purchase order transactions. These cases include 71,811 events executed by 140 employees. Modelling the process flow reveals 5 activity types. The activity types include Create Purchase Order, Signature, Goods Receipt, Invoice Receipt, and Release. Figure 9 shows the modelled process flow for the procurement business process based on the event log. Note that the number under each activity relates to the number of times the activity happened, while the number next to each arrow indicates how many times the transaction flow included going from one activity to the next.

Figure 9. Ideal Process Model From P2P Event Log



This model provides auditors with the most frequent way transactions are being conducted in the procurement cycle. Each transaction starts with Create PO activity. The PO has to be properly authorized, which is represented by a Sign activity. This is followed by receiving the goods and the related invoice, which is recorded in the information system as GR and IR respectively. Finally, a Release activity is performed to release the PO to the accounts payable department for payment.

Once the process flow is modeled and the standard process is identified, the next step is to discover different ways transactions are executed in the procurement business process and highlight deviations from the ideal process. Deviations are examined further to determine if a deviation is immaterial or an actual breakdown in the internal control system. For example, the repetition of GR and IR in a given process instance is a deviation from the standard process flow, but could also represent goods being received over time on installments. This is considered an acceptable deviation and not a violation of controls. However, if an IR activity exists without GR then this is an actual breakdown in the internal control system.

Process discovery can be performed using visualization by examining the expanded process map that includes a 100% of paths and activities, such as the one showed in figure 6. Auditors can instantly notice deviations from the standard model by following the different paths taken by process instances. For example, the expanded process map shows that 4 POs were released for payment prior to receiving the invoice from the supplier.

Hence, conducting process discovery analysis reveals a large number of variants. This is typical to any population data analysis and the result of necessities in operating the business. The P2P process had 1,061 unique sequence of activities in the event log and 4,270 cases. However, only 2 variants account for over 41% of the total number of transactions (cases). Table 1 shows the 7 most frequent variants.

Table 1. Most Frequent Variants

Variant #	Sequence	Variant Frequency		Cum. Total	Throughput Time		
			%	%	Mean	Min	Max
1	Create PO —> Sign —> GR —> IR —> Release	927	21.71	21.71	21.3 d	1.5 h	1.08 y
2	Create PO —> Sign —> Sign —> GR —> IR —> Release	860	20.14	41.85	29.2 d	0.3 h	1.13 y
3	Create PO —> Sign —> GR —> Release	160	3.75	45.60	27.0 d	4.8 h	0.98 y
4	Create PO —> Sign —> Sign —> GR —> Release	115	2.69	48.29	32.1 d	17.9 h	1.02 y
5	Create PO —> Create PO —> Sign —> Sign —> GR —> GR —> IR —> IR —> Release —> Release	101	2.37	50.66	15.5 d	16.0 h	0.77 y
6	Create PO —> Create PO —> Sign —> Sign —> Sign —> Sign—> GR —> GR —> IR —> IR —> Release —> Release	70	1.64	52.30	24.2 d	1.1 h	0.79 y
7	Create PO —> Sign —> Sign —> GR —> IR —> GR —> IR —> Release	50	1.17	53.47	199.5 d	9.1 d	1.13 y

From the initial analysis of the most frequent variants found the P2P event log dataset, it is noticed that there is repetition in some tasks (events) in the transaction sequence. For example, the only difference between the top two most frequent variants from an event perspective is the addition of a *Sign* activity. This is due to a business rule that requires any PO created with a value above a certain threshold ($5,000, as indicated by the business rules of the P2P process) to have two signatures by two authorized employees. However, when it comes to the third and fourth most frequent variants, there is a breakdown in the purchasing process as these transactions are missing a key control, which is an *IR*. This sequence of *PO*s are being released and billed without a valid invoice entered in the system, which violates a business rule. As for variants five and six, each case represents two different *PO*s created for different values. The event log recorded these *PO*s as one case since they were created with the same timestamp by the same employee. These cases are considered to be normal since they follow the verified business rules. The final variant shown in Table 3 is also considered a normal business activity from an event sequence perspective since goods and invoices are being received on installments over a period of time.

This large number of variants found in the P2P process can overwhelm auditors if they do not have sound guidance to utilize this information. There needs to be a baseline in the form of a set of rules that can compare these variants against to determine whether the variants are acceptable deviations from the designed business process model that had to be overridden for operational necessity, or they are a violation of control procedures. The business rules are checked using conformance checking technique to determine where the violation is occurring.

Additionally, notice that the observations found in Table 1 only relate to the sequence of events analyzed for an event log. It does not consider other aspects such as the event initiator (resource), the financial value of the transaction, or even the duration of time it took to complete the transaction. Hence, a rule-based process mining technique is essential and proves to be valuable in uncovering deficiencies in the internal control system.

To able auditors in determining if the different variants and process deviations are indicative of control violations, risks and controls need to be identified. Based on the process discovery step, and risks associated with the P2P business process, a set of risks are identified. Table 2 provides a list of the risks identified.

Table 2. Relevant Purchasing Process Risks

Risk	Description
Inappropriate purchase order	Purchase order is inappropriate because: - the purchase order does not match a valid requisition; - purchase is at the incorrect price; or- an inappropriate or unauthorized vendor is selected.
Incorrect invoice approval	Invoice approved for payment at incorrect price or at incorrect quantity of goods/services received or before services received
Inappropriate access	An employee may have the ability to initiate, authorize and record a transaction or may have custody of assets within the process, such that they are able both to perpetrate and conceal an error or irregularity.

Goods received not matching valid purchase	Goods received did not match a valid purchase order.
Payment error	payments are not booked, recorded to the incorrect vendor account, recorded for an incorrect amount, or made for goods/ services not received.
Suboptimal task allocation	purchase processing costs might significantly rise when activities are performed by overly qualified employees or when there are unnecessary hand-overs.

The next step is for auditors to link the identified risks with mitigating controls. Controls can me mapped as rules (Caron et al. 2013). Rule-based controls allow auditors to test the effectiveness of controls based on any deviation from the ideal process, while being able to differentiate between an acceptable deviation and an actual breakdown in the internal control system. Therefore, the rules created in this step consider all the risks associated with the business process and the controls that need to be in place to mitigate the risks. Table 3 provides a list of the rules and the linked risks. It needs to be noted that the overall objective in the audit of the P2P process and its most important assertions relate to the evaluation whether the acquisitions of goods and services and the cash disbursements for those acquisitions are fairly presented in the accounts in accordance with the generally accepted accounting principles (Arens et al. 2003).

Table 3. Purchasing Process Rules & Controls

Risk	Assertion	Control
Inappropriate purchase order	Existence	A *sign* activity must be performed at least once
Inappropriate purchase order	Existence – Valuation	The value of a *purchase order* must be specified
Inappropriate purchase order	Existence – Valuation	The value of a *purchase order* may not change after a *sign* activity has been performed
Incorrect invoice approval	Existence - Timing	A *purchase order* activity must be started before date of *invoice receipt*
Inappropriate access	Occurrence – Existence – Valuation	A person must not perform all activities of the P2P process
Inappropriate access	Timing	A *good receipt* activity must be performed during regular business hours
Inappropriate access	Timing	A person must perform a *release* activity after time T = timestamp of *goods receipt* event
Inappropriate access	Occurrence – Existence – Valuation	A *release* activity must be performed by a member of *senior staff*
Invoice entry error	Existence – Accuracy	An *invoice pay* activity cannot be duplicated for the same *purchase order*
Goods received not matching valid purchase	Existence – Accuracy – Posting & Summarization	The values of *Purchase order*, *goods receipt*, and *invoice receipt* must match before the corresponding invoice can be paid

Payment error	Existence – Posting & Summarization	If a *goods receipt* activity is performed then an *invoice receipt* activity must be performed
Suboptimal task allocation	Classification	A *good receipt* activity must not be performed by a member of *senior staff*

2.4.3. ASSESSING THE OPERATIONAL EFFECTIVENESS OF THE INTERNAL CONTROL SYSTEM

After assessing the design of the internal control system, which allows auditors to determine whether controls are implemented or not, the next component of the framework is assessing the effectiveness of the implemented controls. This component examines the operational effectiveness of the implemented controls.

Traditionally, auditors assess the operational effectiveness of controls by sampling from a large pool of transactions and determine if the sample followed the process and control design. And if a deviation occurs, then a follow-up is needed to determine the cause of the deviation.

However, this approach only provides a very limited scope on the "real" effectiveness of controls. The only way to determine the overall effectiveness of controls is by testing every instance of it, which is usually accomplished by using analytical techniques. This framework demonstrates how process mining is a very effective tool in evaluating control risk and provide strong evidence on their operational effectiveness.

The analysis is done by examining meta-data timestamps to systematically establish the flow of activities for each PO line, from creation to payment. This type of analysis is

unique to process mining because it utilizes process-related data instead of only static transactional data. Using traditional analysis techniques would not yield these insights because they rely only on data entered by the auditee that cannot be compared with independent system information (Jans et al. 2014). Additionally, when testing the effectiveness of controls, auditors need to consider the impact the failed controls have on the financial statement. This helps in assessing the control risk and its place on the risk map.

The procurement process event log included 4,270 cases with a total PO value of $95,821,927.49. The total number of cases or process instances were used to calculate the violation frequency to determine the EoC, while the total PO value was used to calculate the severity or the impact those violations have on the financial statements.

The first controls tested relate to the first management assertion risk in the procurement process risks and controls model, which is inappropriate purchase order. Here, auditors are testing if a PO is inappropriate due to not matching a valid requisition, at the incorrect price, or an inappropriate or unauthorized vendor is selected. This risk is mitigated by having every PO signed by an authorized employee that checks for all these risks. The result of testing these controls found 0 cases that did not have a Sign activity for the PO nor any changes made after the Sign activity, while 69 cases lacking any value for the PO. This gives the authorization of PO an EoC of 100%, and PO value control an EoC of 99%, which makes it an Effective Internal Control. However, the severity of the POs without a specified value equaled to $1,251,516.22. Auditors can consider that control to have a material impact on the financial statement even though it is effective from an operational perspective. Therefore, this framework provides auditors with direct evidence

on the level of control risk to allow them to objectively issue an opinion on the effectiveness of controls.

The second risk in the procurement process risks and controls model is incorrect invoice approval. Here, the auditors test for the risk that invoices are approved for payment at incorrect prices or at incorrect quantities of goods received. The control that is in place to mitigate that risk is prevention of an invoice being received before the creation of a PO. There were 6 cases that violated this control with a transactional value $241,594.87. This controls appears to be operating effectively and has an EoC of 99.9%.

The third risk in the procurement process risks and controls model is inappropriate access. In this case, auditors consider the risk that an employee may have the ability to initiate, authorize and record a transaction or may have custody of assets within the process, such that they are able both to perpetrate and conceal an error or irregularity. To mitigate this risk, management has in place several different set of controls: segregation of duties, receipt of goods during regular business hours, release and payment of PO has to be after goods are received, and the release of the PO has to be by senior staff. Segregation of duties is a crucial control that needs to be implemented throughout the purchasing process, not only a subpart of the process. Therefore, process mining has revealed that the only violation of segregation of duties in the whole population of data involved 16 cases where the same employee performed all activities in the P2P process. The violated PO had an amount of $42,853.57. Even though this control can be considered effective from a frequency and impact perspectives, auditors might consider this control to be ineffective on the basis that there shouldn’t be any possibility that the same employee is able to perform all activities in the P2P process. This reflects on the design of the internal control system and increase

control risk. As for the other three controls in place to mitigate the risk of inappropriate access, the verification of their existence correctness and functioning show that they have an EoC of 99%, 97%, and 99% respectively. However, the release and payment of PO has to be after goods are received control failed in 32 cases, they amounted to $10,584,025, which can be considered a material amount since it is equal to 10% of the total PO amount. The combination of EoC and impact can affect the final risk assessment for the internal control system.

The fourth risk is the goods received not matching valid purchase. Here, the auditors test for the risk that invoices are approved for payment at incorrect prices or at incorrect quantities of goods received. The control that is in place to mitigate that risk is 3-way match. The control is in place to assure that for every PO created, an IR and GR are available and matched before payment. Out of the 4,270 POs in the event log, 1,310 cases have violated this rule, which account for 31% of total POs created. These exceptions could be rationalized by legit business operations, where exceptions could be normal and non-fraudulent. However, allowing exceptions for the ideal operational design of the business should be considered when evaluating the internal control system for organizations, and hence, will affect the score of such controls for different business processes. Therefore, the 3-way match control has an EoC score of 69% indicating that it is a Materially Weak Control. Additionally, the total monetary value for the violated POs is $32,932,037.54. This is a very large amount and accounts for 33% of the total POs. The testing for this control might be adjusted using some of the business rules from the organization. However, since such business rules are not available, the results couldn't be adjusted.

The fifth risk is payment error risk, which is the risk that payments are not booked,

recorded to the incorrect vendor account, recorded for an incorrect amount, or made for goods not received. The control in place is that a GR activity must be accompanied by an accurate IR activity. The results of testing this control show that 480 POs had goods received without an invoice from the supplier, indicating an Significantly Deficient Internal Control. In addition, the failed control had an impact of $7,451,027.95, which is material. This increases the risk assessment for this control. Auditors can consider this risk to be materially weak since it amounted for a material amount.

The last risk assessed is suboptimal task allocation. Auditors might consider this risk to be low since it does not affect the financial statement directly due to its focus on resource optimization. But, process mining allows auditors to consider different types of risks to provide stronger evidence of the controls implemented in the business process. This risk is mitigated by the control that certain tasks should not be completed by certain roles. In the P2P process, good receipt activity must not be performed by a member of senior staff. 36 cases violated this control. It has an EoC of 99% indicating and Effective Internal Control. Table 4 shows the results of the procurement process risks and controls model.

Table 4. Evaluation of The Procurement Process Risks and Controls Model

Risk	Control	Violation Frequency	%	Severity	%
Inappropriate purchase order	A *sign* activity must be performed at least once	0	0%	$0	0%
Inappropriate purchase order	The value of a *purchase order* must be specified	69	1%	$1,251,516.22	1%
Inappropriate purchase order	The value of a *purchase order* may not change after a *sign* activity has been performed	0	0%	$0	0%
Incorrect invoice approval	A *purchase order* activity must be started before date of *invoice receipt*	6	0.1%	$241,594.87	0.3%
Inappropriate access	A person must not perform all activities of the P2P process	16	0.4%	$42,853.57	0.4%
Inappropriate access	A *good receipt* activity must be performed during regular business hours	61	1%	$989,351.47	1%
Inappropriate access	A person must perform a *release* activity after time T = timestamp of *goods receipt* event	139	3%	$10,584,025.10	10%
Inappropriate access	A *release* activity must be performed by a member of *senior staff*	27	1%	$683,955.64	1%

Invoice entry error	An *invoice pay* activity cannot be duplicated for the same *purchase order*	32	0.7%	$844,426.38	0.5%
Goods received not matching valid purchase	The values of *Purchase order*, *goods receipt*, and *invoice receipt* must match before the corresponding invoice can be paid	1310	31%	$32,932,037.54	33%
Payment error	The value of a *purchase order* may not change after a *sign* activity has been performed	0	0%	$0	0%
Payment error	If a *goods receipt* activity is performed then an *invoice receipt* activity must be performed	480	11%	$7,451,027.95	7%
Suboptimal task allocation	A *good receipt* activity must not be performed by a member of *senior staff*	36	1%	$763,489.34	1%

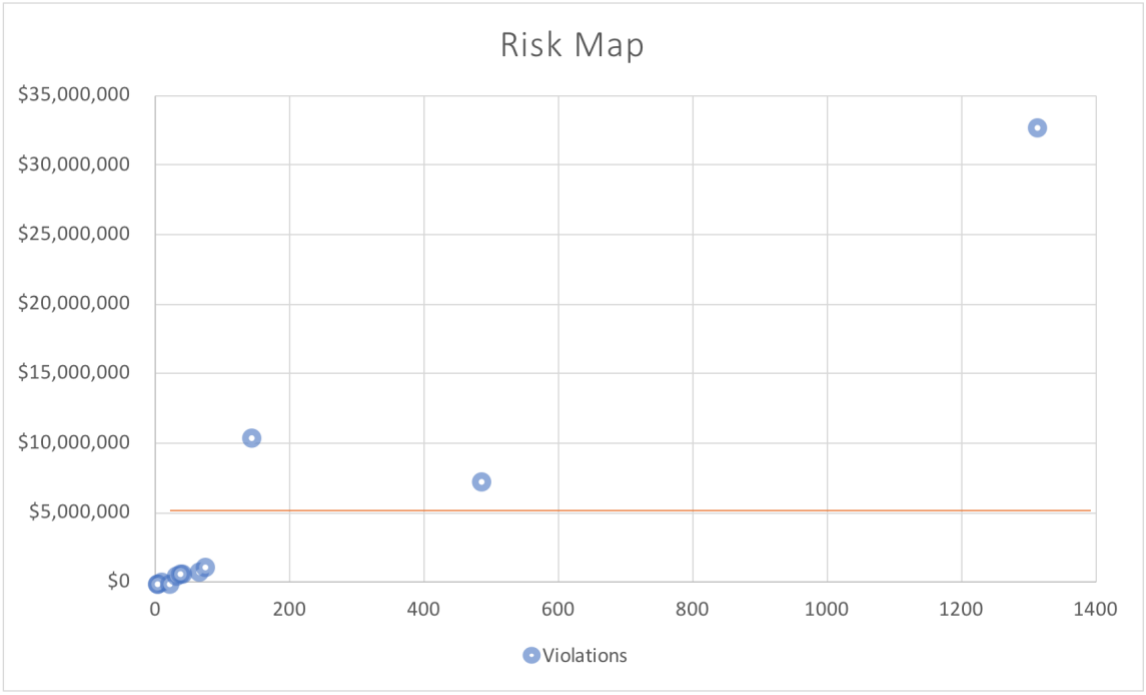
It should be noted that there are concerns with calculating the effectiveness of the internal control system for a business process. One issue is concerned with controls that are not included in the testing or the scores calculated. These controls would include manual, but essential, controls. For example, having periodical reviews of transactions, explanations and invoices; or doing physical examination of goods received. Some of these controls could be formalized and reflected in the information system. Another concern would be in the quality, detail, and accuracy of the data being tested. Any results obtained from testing the data depends mainly on the mentioned variables. The more detailed the

data, the more tests can be done on it and more results obtained.

2.4.3. RISK ASSESSMENT

The results of testing the effectiveness of controls are used by the auditor for control risk assessment and making a decision on the risk appetite. Risk appetite is defined as the level of risk and uncertainty auditors are prepared to accept (Caron et al. 2013). Risk appetite can be represented graphically as a risk map and is a function of violation frequency and monetary impact of deviated transactions. Figure 10 illustrates the risk map for the P2P process for the identified risks and controls and ranked according to their frequency and impact.

Figure 10. P2P Risk Map



Visualizing the violations and their impact provides auditors with an instant way of highlighting areas of increased risk by viewing anomalies on the risk map. For example, the 3-way match control is an obvious high-risk control since it is way above the materiality line in terms of impact and has the most frequent violations. Auditors can opt to focus on testing in more detail the controls that are associated with risks above the materiality line on the risk map. The materiality line is determined based on what the auditor deems to be a material amount in comparison to the whole financial statement. Alternatively, auditors can focus frequency of violations for control effectiveness testing. The risk map assists auditors in control risk assessment by utilizing visualization. The main objective of implanting a graphical representation of control risk is to help auditors gain better insights, draw better conclusions, and assess risk objectively.

2.5. CONCLUSION

This paper described a methodology for objectively measuring the effectiveness of internal controls and risk assessment. Instead of relying on traditional and qualitative methods, the general framework in this study would provide auditors with a more objective and efficient way of assessing if controls are implemented, and to what degree. The general model was applied to a P2P business process from a national non-profit organization.

The controls tested in this study were based on industry standards and literature, not entirely from what the firm that the data is generated from has in place. Therefore, this is a limitation of this study since some controls were simulated. It would have been more preferable if the controls that the firm had in place were known so that they could be

measured in terms of their operating effectiveness. Although, it should be noted that by using industry best practices and practitioners to aid in understanding which controls should be in place, this step is not unlike the information gathering and subsequent brainstorming that auditors undertake, and should mimic these steps as much as possible.

This paper proposed a measurement approach for evaluating the effectiveness of internal controls that could be tailored to different industries and business processes. As such, it contributes to the sparse literature on internal controls effectiveness measurement. In future work, one can utilize different weighting schemas and various variables to enhance the calculation method and provide a more accurate measure for risk assessment.

CHAPTER 3: PROCESS INSTANCES RISK PRIORITIZATION

3.1. INTRODUCTION

Complex ERP systems capture thousands if not millions of transactions on a daily basis, and with this large amount of data, it is impractical to analyze it using traditional and periodic techniques. Many organizations, after implementing ERP systems, are still depending on manual procedures. Due to this, fraud may go undetected for extended periods of time. Hence the use of advanced audit analytics on a daily continuous basis is necessary to detect, and possibly prevent fraud.

However, despite the use of such advanced techniques, an overwhelming amount of exceptions are generated (Alles et al. 2006, 2008; Debreceny et al. 2003), causing the overall efficiency to decrease due to the limitations of human processing. Alles et al. (2006, 2008) and Debreceny et al. (2003) discussed this issue and pointed out that these exceptions are generally generated and sent to auditors without prior processing or sub-filtering. These scenarios raise the question of how users can organize and make sense of such voluminous data.

With the large number of transactions being executed on a daily basis, auditors are facing more difficult ways at detecting and investigating anomalies and exceptions. Issa (2013) attempted to resolve the issue of information overload by proposing methodologies that would prioritize exceptions. Such attempts can help auditors focus on the more suspicious cases and make further investigation be more efficient. This paper proposes a methodology where process mining can be used to apply a suspicion function for each

transaction to assign it a risk score. The risk score is based on different criteria, such as the total number of violations the transaction commits, the severity of the violations, and the monetary amount of the transaction. The methodology implemented in the study will allow auditors to objectively determine the riskiness of transactions. Additionally, from all the anomalies or suspicious transactions found, the methodology would guide auditors to the riskiest transactions that require further investigation, while filtering out low risk transactions.

Process-aware information systems are "software system that manage and execute operational processes that involve people, applications, and/or information sources on the basis of process models" (van der Aalst 2009). These systems, such as ERPs, allow for dynamic process and service changes. This, in turn, has led to one of the main challenges for process mining, which is the large number of process model variants. This large number of process variants are difficult to maintain and expensive to configure (Li et al. 2008). Companies allow for flexibility in their business processes and will inevitably incur process variances to allow for exceptional transactions. For example, three-way match may not always be realized due to the inclusion of unanticipated transportation costs that were not included in the original purchase order. In theory, the ERP system can be configured in such a way that such deviations are not allowed and, hence, become impossible to execute. But locking down the process in this way would result in a constant stream of exceptions and delays since the actual procurement cycle would often deviate from the designed process for a variety of reasons, some anticipated and acceptable, and some not. For example, there could be problems in manufacturing the items.

The introduction of process mining techniques in the past decade is very promising from an auditing perspective. It allows for population testing on the entire set of real process executions that is automatically recorded in the information system and compare the event log with the ideal process model to identify deviating transactions from normal ones (Rozinat and van der Aalst 2008; Adriansyah et al. 2011). However, process mining being used in auditing is not without its shortcomings. For example, conformance checking technique (comparing real process instances with a process model for conformance) can result in a large number of detected deviations, which can be too immense for auditors to follow-up on (Hosseinpour and Jans 2016). The large number of variants is a result of normal business operation where it must allow for flexibility in executing processes to accommodate customers' needs.

The motivation behind this study is that when auditors resolve to population testing instead of the traditional option of taking a sample, it results into a large number of anomalies or exceptions that can overwhelm the auditors. Even if auditors want to investigate all those anomalies, it might be impossible to do so due to the large number. This problem is inherent to population testing. Therefore, providing auditors with a framework that comprises of multiple stages of filtering and a suspicion function for prioritization based on the riskiness level of each transaction to determine the ones that are most likely to be highly problematic would be very beneficial for auditors to avoid or minimize the downside of population testing. This solution would allow auditors to validate whether the audit-relevant information generated by process mining is really indicative of fraud, while avoiding having to deal with a flood of false positives that would arise when any analytical procedures are applied to the entire population of data.

This paper contributes to the auditing literature by proposing a methodology that provides auditors with guidance as to which notable transactions need further investigation. The identification and prioritization of such risky process instances helps with the information overload problem that entails population testing. In addition, this paper attempts to provide a solution to one of the challenges auditors face when applying process mining in their audit engagement, which is the large number of false positive variants. Also, this paper provides guidance on the use of process mining in conjunction with existing analytical procedures to allow auditors to focus on process instances that are likely to be considered high-risk, reduce the risk of failing to detect material misstatement, and enhance audit effectiveness.

The framework proposed in this study was demonstrated on a real-life event log dataset that was obtained from the procure-to-pay process of a not-for-profit national organization. The event log contained a total of 4,142 process instances. After applying the first part of the framework, which is the process mining part, it highlighted 1,346 notable process instances. Existing analytical procedures were then applied, which is the second part of the framework, this resulted in narrowing down the results of problematic process instances to 814 exceptional process instances. A threshold was then applied to focus on process instances with a monetary value above a certain amount, which resulted in highlighting only 457 highly problematic process instances that have a material amount.

The remainder of the paper is organized as follows: section (2) provides a background and literature review on employing analytics in auditing and population testing. This will be followed by section (3), which describes the methodology and general framework developed in this study. Section (4) provides an illustration of the methodology

on a specific business process, and describe the data used, the analysis, results, and discussion. Finally, section (5) concludes the paper.

3.2 BACKGROUND

As technology has become the norm in operating businesses and an increasing number of companies have implemented information systems, it is apparent that auditors must take advantage of the availability of big data in the different stages of the audit. However, it is well documented that the emergence of big data as well as the increase in adopting data analytics in business processes has brought new challenges to the audit community (Vasarhelyi et al. 2015; Appelbaum et al. 2017). Some of these issues and challenges are concerned with what type of analytics is most appropriate and in which part of the audit are they suitable (Appelbaum et al. 2017). One of the most promising analytical tools available to auditors is process mining. The main objective of process mining is ‘‘to discover, monitor and improve real processes (i.e., not assumed processes) by extracting knowledge from event logs readily available in today’s systems’’ (Van der Aalst 2011).

One of the side-effects of organizations’ ongoing automation of business process is the unused process data that is available, which can be used for process mining (Azzini and Damiani 2015). The automation of business processes leads to having digital traces of real process executions. These digital traces reflect what is actually happening in the real world and enable the application of process mining (Azzini and Damiani 2015). By applying process mining, organizations can understand how their processes are actually executed and eventually gain control over their complex business environment. Auditors can use the

same information to find evidence on the effectiveness of controls and whether financial statements are clear of material misstatements (Van der Aalst et al. 2010; Jans et al. 2011, 2013, 2014).

The auditing profession in the United States is overseen by the Public Company Accounting Oversight Board (PCAOB). Auditing standards issued by the PCAOB emphasize the importance of understanding the processes that make up the financial statements when doing the audit (Jans 2011). According to Auditing Standard No. 5 (paragraph 34), auditors must understand the flow of transactions and identify the controls implemented by management to address potential misstatements or to prevent unauthorized acquisition, use, or disposition of a company's assets. The traditional way of complying with that standard, and what is actually recommended by it, is by doing walkthroughs. Standards consider the use of walkthroughs as the most effective means to understand processes. The way walkthroughs are performed is by following the path of a transaction as it flows through the different steps in a business process form initiation to completion and reflected in the company's financial records. This manual approach, which is currently used in the auditing profession, can be significantly improved by employing process mining techniques (Jans 2011). Process mining can not only automate walkthroughs, but also extend the analysis to the full population instead of a sample. This results in a transparent overview of the process (Jans 2011).

Understanding the process is one part of the audit that process mining can excel over other methods. However, process mining techniques' strengths are in process centric analysis. Therefore, Damiani and van der Aalst (2015) have argued that there needs to be a careful combination of process-centric and data-centric approach to analyzing business

processes. The data-centric analysis that can be accomplished by using other existing analytical procedures. SAS 56 (American Institute of Certified Public Accountants 1988) requires that analytical procedures be performed during different stages of the audit, such the planning and review stages. The standard also recommends their use in substantive testing so that auditors limit the subsequent test of details to areas that constitute high risk and concern (Kogan et al. 2014). Auditing Standard No. 15 defines analytical procedures as the "evaluations of financial information made by a study of plausible relationships among both financial and nonfinancial data" (PCAOB 2010, paragraph 21). Since analytical procedures are required in different phases of the audit, so does its purpose (Appelbaum et al. 2017). For example, analytical procedures in the risk assessment/planning phase should enhance the auditor's understanding of the business and its transactions and highlight areas that might be problematic and high-risk to the audit.

Applying process mining and other analytical procedures in different stages of the audit results in identifying anomalies and areas of concern. Having a prioritization method for the identified suspicious transactions and anomalies can significantly benefit auditors and minimize the effects of information overload, especially in the risk assessment stage of the audit (Kim and Vasarhelyi 2012; Issa and Kogan 2014; Li et al. 2016). Kim and Vasarhelyi (2012) argued that using the knowledge engineering of experienced professionals (Vasarhelyi and Halper 1991) allows for auditors to determine risk factors or indicators of abnormality that could be considered problematic and fraudulent in nature. These indicators were weighted by giving each a score based on risk. Anomalies were then prioritized based on the risk score of each. Issa and Kogan (2014) argue that processing and prioritizing the large number of outliers identified in population testing can help

auditors overcome the human limitations of dealing with information overload and consequently improving overall audit efficiency by focusing on the most suspicious transactions. Li et al. (2016) propose a framework for prioritizing exceptional transactions based on the likelihood of it being erroneous or fraudulent. They proposed this frameowk as a solution for the negitive effects of infromatoin overload, since they found that the volume of exceptions generated by a continuous auditing system can be overwhelming for auditors to handle

3.3. METHODOLOGY

The methodology that is used to identify process instances that are deemed to be most risky is based on the application of process mining to an event log in conjunction with existing analytical procedures. A set of filters based on business rules mapped to the risks and controls for the targeted business process. Each rule is given a weight based on importance and relevance using auditor judgment. Then, depending on the number of violations for each process instance along with its monetary value, a risk score will be given. The risk score (depending on the number of rules violated) in addition to the monetary value of it (above or below a certain threshold) with the tested violations will result into the ranking of violated process instances.

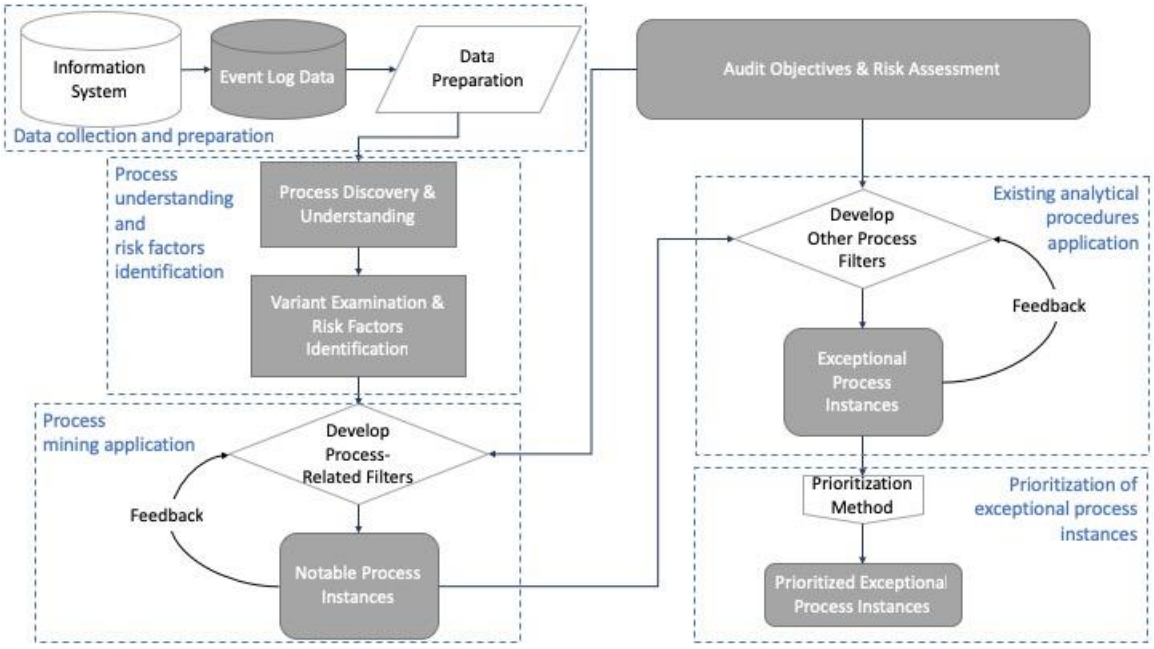
The methodology of this study expands on the process mining risk assessment framework proposed in Chiu et al. (2018)'s study. This study provides guidance on how process mining can be used in the audit process in conjunction with other analytical procedures and tests that cannot be done relying solely on process mining analysis. The methodology will be illustrated using a real-life dataset. The results of the demonstration

will be compared to those of Chiu et al. (2018) study. This methodology can also be used as basis for implementing a continuous process monitoring system, similar to Alrefai (2019)'s study. For example, the methodology can be used to determine whether a tested transaction will be allowed or blocked based on its risk score.

3.3.1. FRAMEWORK

The process instances risk prioritization framework is comprised of five stages: 1) Data collection and preparation, 2) process understanding and risk factors identification, 3) process mining application, 4) existing analytical procedures application, 5) prioritization of exceptional process instances. The final three stages of the framework are implemented with the audit objectives and auditors' risk assessment of the underlying process in mind. The overall flow of the framework is found in figure 11.

Figure 11. Process Instances Risk Prioritization Framework



Data Collection and Preparation

Before any analysis can be done, there needs to be a set of data that auditors can apply their analysis to. In the case of process mining analysis, specific type of data needs to be collected and transformed. This set of data is in the form of an event log. Every event log contains three main types of information, and without it, the event log would be considered insufficient for process mining analysis. The three main types of information are: activity, resource, and timestamp. The activity describes what step of the process has been performed for a specific process instance. The resource provides the name of the user who performed the activity. The timestamp provides the date and time of when the activity was performed. Table 1 is an example of an event log template for a procure to pay process.

Table 5. Event Log Template

CASE ID	ACTIVITY	NAME	TIME STAMP	VALUE PO	QUANT. PO	QUANT. GR	VALUE PAY	VALUE GR	SUPPLIER
87127	Create PO	Paul	8/3/15 9:57	25000	25,000				AC PROS
87127	Sign	Paul	8/3/15 9:57	25000	25,000				AC PROS
87127	Sign	Tiffany	8/7/15 14:18	25000	25,000				AC PROS
87127	IR	Beverly	8/11/15 12:02	25000	25,000	25,000	25000	25000	AC PROS
87127	GR	Beverly	9/1/15 10:50	25000	25,000	25,000		25000	AC PROS
87127	Release	Kimberly	9/1/15 10:53	25000	25,000				AC PROS
87128	Create PO	Paul	8/3/15 10:58	21250	21,250				HS INC
87128	Sign	Paul	8/3/15 10:58	21250	21,250				HS INC
87128	Sign	Tiffany	8/7/15 9:54	21250	21,250				HS INC
87128	IR	Beverly	8/10/15 10:13	21250	21,250	21,250	21250	21250	HS INC
87128	GR	Beverly	8/20/15 10:35	21250	21,250	21,250		21250	HS INC
87128	Release	Jay	8/20/15 12:35	21250	21,250				HS INC

The event log in Table 5 shows other information besides the typical information found in an event log. In this case, Table 5 contains financial and non-financial data that are valuable from an auditing perspective. For example, the event log template shows the dollar amount of each purchase order (PO) along with the value of the goods receipt (GR) and amount payed for any invoice receipt (IR). This additional information allows for supplementary analysis to process mining using existing analytical procedures that is key to the process instances risk prioritization framework.

The detailed event log data required for the implantation of the framework is found in the information system of the organization. Enterprise Resource Planning (ERP) systems, such as SAP, are capable of creating event logs for different business processes.

However, companies tend to switch off the logging capability since it taxes the system both in terms of performance and storage. Therefore, event logs are not created automatically. The event log that of the intended business process must be developed using knowledge about the ERP system and its table structure for the underlying business process.

Once the event log is created and the necessary fields are included, the next step in this stage is data preparation. Data preparation is essential to any analytical technique and ensures that the data can analyzed with the least amount of noise found in it. For example, any process instance that is incomplete can be removed from the dataset in this step, so that the results are not biased or magnified.

Process Understanding and Risk Factors Identification

The second stage of the framework involves one of the most important insights that process mining provides to auditors, which is process discovery. To complete this stage, the event log dataset is imported into a process mining application (i.e. Disco or ProM) and examined for both process understanding and risk factors identification from non-standard variants. Process discovery shows auditors how the process is actually operated in the organization. Management provides auditors with how a standard transaction is conducted in a specified business process, but process discovery shows auditors whether the standard method prescribed takes place and how frequent it does. This allows auditors to know what is actually happening throughout a business process by examining every instance of it. Numerous deviations from the designed process model are often found in a given business process to ensure smooth operation of the business. Therefore, auditors usually find a large number of variants when analyzing a business process and need to be able to determine of which is considered acceptable and which would not be.

The second step of this stage is variant examination and the identification of risk factors. Auditors can start this step by first examining the different variants found and identifying risks associated with non-standard variants. Auditors can identify anomalies from the non-standard variants by performing the process discovery step. In addition, auditors can rely on their prior experience since there are usually general risks associated with a given business process.

Process Mining Application

The third stage of the process instances risk prioritization framework is the process mining component of it. In this stage, auditors first develop a set of process-related filters that focus on identifying the attributes of high-risk notable process variants. These filters are based on the "What Could Go Wrong" (WCGW) and risk associated with the underlying business process. Auditors should also consider their prior experience when developing these process-related filters. The process-related filters should be in line with the objective of the audit that need to be achieved during test of controls and substantive test of details. The filters should encompass controls that need to be in place to mitigate the risk of management assertions.

Once the filters are developed, auditors apply them to the entire population to discover high-risk notable process instances. After initial results, auditors can evaluate whether the design and performance of the filters are acceptable. Auditors can either modify or confirm the final set of process-related filters and obtain the notable process instances as a subset of the entire population. Notable process instances are identified by either failing one or multiple filters.

Existing Analytical Procedures Application

Having a methodology that provides guidance to auditors on the integration of process mining with other analytical procedures is unique to this study. In this stage of the framework, auditors apply existing non-process mining related analytical procedures to the notable process instances in order to filter out exceptional outliers that would be considered most risky and highly problematic. This second step of filtering reduces the number of outliers found in the process mining application stage. For example, a process mining filter can examine process instances that had multiple GR activities for the same PO. From a process mining perspective, this shows a deviation from the ideal process model and might indicate excess shipments from the supplier. However, implementing other existing analytical procedures, like a 2-way match between PO value and GR value, can reduce the number of notable items and filter out noise found in those results.

Filters applied in this stage are different from those applied in the process mining stage, but they complement one another. This is due to the inherent limitation of process mining where its strength lies in understanding the actual process and discovering anomalies and deficiencies that relate to how processes are executed and the steps taken. However, other procedures and controls cannot be tested or is very difficult to do so using process mining techniques. For example, if an auditor wants to test if a series of purchase orders with same employee-vendor match is splitting purchases. This test cannot be achieved using process mining techniques. Therefore, auditors need to use existing analytical procedures in conjunction with process mining to have a more effective and efficient audit. However, in order to do so, a detailed event log with financial and other information is required.

The existing analytical procedures used in this stage of the framework are based on the risks and WCGWs associated with the underlying business process in which process mining filters did not test. The purpose of adding this stage of filtering to traditional process mining analysis is to provide auditors with guidance to the riskiest sub-sample of the entire population of process instances in which they can either further examine or sample from, depending on the final number of exceptional process instances. The main reasoning behind this additional stage is that if a process instance violates one or more process mining filter, and therefore is considered a notable process instance, and it violates a second set of high-risk filters that is not mainly concerned with its routing, then that would indicate that the process instance is highly problematic and therefore exceptional.

In this stage, auditors can implement a materiality cut-off if the number of exceptional process instances is large to work as a third filtering method for high-risk process instances. In this framework's context, a materiality cut-off can be viewed as any process instance that has a dollar amount that an auditor would consider material (i.e. more than $5,000). Auditors can also consider a materiality cut-off if their judgment and risk assessment urges them to examine material process instances only.

Prioritization of Exceptional Process Instances

The final stage of the framework is prioritizing the exceptional process instances by applying a risk score to each one and ranking them. The risk score formula is adapted from Issa (2013) and can be calculated as:

$$Risk\ Score\ (X_i) = \sum W_{Fj}\ V_{Fj}\ M_{X_i},$$

Where

X_i *refers to the exceptional process instance,*

W *is the risk level of filter* F_j*,*

V *is the binary variable that euals* 1 *if process instance* X_i *violates filter* F_j*,*

and 0 *otherwis,*

M *is the monetary value of process instance* X_i

The risk score for each exceptional process instance considers the risk level of the violated filters. So, as auditors are developing filters for process mining and existing analytical procedures, they need to assign a risk level to each one based on their professional judgment. For example, if a process instance is missing a Release activity, then that filter would be given a level of risk that is high. Risk levels would be based on a high-medium-low scale, based on standard auditing procedure. These levels can be numerically converted into 3-2-1, where 3 is for high risk level filters, 2 for medium risk level filters, and 1 for low risk level filters.

Once the exceptional process instances are prioritized, auditors can either follow up on all of the prioritized exceptional process instances or only a subset of them. This decision is based on the results of stage 4 of the framework. If the number of exceptional process instances is too large for auditors to investigate, then in this case, auditors can choose to investigate only 50 exceptional process instances with highest risk score, for example. Auditors can also choose the exceptional process instances with the highest monetary value if materiality is the highest priority. The decision of determining how many exceptional process instances to follow up on, and which prioritization method to use is

based on auditors' professional judgment, the determined risk of the underlying business process, and the acceptable level of risk.

3.4. ILLUSTRATION OF METHODOLOGY

This section provides an illustration of the process instances risk prioritization framework. A procure-to-pay (P2P) process event log dataset is used for that purpose. The different stages of the framework will be explained and demonstrated to show the usefulness of such framework for practitioners and researchers. However, the illustration of the methodology will be grouped into three main parts: the process mining element, the existing analytical procedures element, and the prioritization element. Concluding this section will be a comparison of the results of this study with Chiu et al. (2018)'s study to evaluate its effectiveness.

3.4.1. PROCESS MINING ELEMENT

Any process mining analysis has to start with an event log dataset. The data used for this study to demonstrate the framework is an event log extracted from the information system of a national not-for-profit organization. Specifically, the event log is from the organization's procurement business process. As explained in the framework, the event log must contain not only process related information such as the activities, users, and timestamps for all process instances in the business process, but also other non-process related information such as quantity and monetary value of each process instance. The P2P event log includes monetary amounts for the value of the purchase order, the goods

received, and the invoice for all process instances. This transactional value is what makes this event log dataset unique and expands the application of process mining that is available to auditors with the availability of such information. Additionally, this event log dataset contains information on the suppliers that were involved in all POs. The structure of the event log is similar to that found in Table 5.

The event log contains a total of 4,270 process instances that started from June 2016 till December 2016. There were 71,811 number of events happening throughout the process. The P2P business process includes 5 main activities: Create PO, Sign, GR, IR, and Release. The process instances in the event log were grouped into 1,061 variants, which means that transactions of the purchasing cycle were executed in 1,061 different ways for the period being analyzed. This shows that employees in practice do not necessarily follow the standard way of conducting transactions in the purchasing cycle, and therefore, violations of business rules and controls might occur. Even with a relatively small event log dataset, a large number of variants is present. This emphasizes the need for guidance on how to filter out notable variants from acceptable ones and complementing those findings with existing analytical procedures to identify highly problematic process instances.

One of the main reasons why this study grouped the first three stages of the framework into one category "process mining elements" is because, in some instances, data preparation, process discovery and understanding, and variant examination and risk factors identification steps can be done simultaneously. When applying process discovery and understanding, this can lead to other data preparation that was not included prior to this step. For example, when first analyzing the data set, there were 128 process instances with

21 variants that lacked both GR and IR activities. These process instances all had a PO value of $0 and were cancelled by the organization. Leaving these process instance in the dataset was causing noise in the overall results of some of the preliminary analysis in the process mining stage, and therefore, needed to be removed. Additionally, days of the week were added to the event log based on the timestamp in order to apply related filters. Table 6 describes the event log dataset used in this study to illustrate the framework after preprocessing.

Table 6. P2P Event Log Descriptive Statistics

Events	71,203
Process Instance	4,142
Number of Activities	5
List of Activities	(1) Create PO (2) Sign (3) GR (4) IR (5) Release
Variants	1,040
Resources	140
Mean Process Instance Duration	79.3 Days
Start	06/10/2014
End	12/02/2016

The frequency of activities in the event log is summarized in Table 7. Note that if the designed procurement process is followed, then all activities should have the same

frequency. However, this is not the case, which provides immediate evidence that the actual purchasing process differs from the designed process, indicating deviations from the standard purchasing process model.

Table 7. Frequency of activities in the event log

Activity	Frequency
Create PO	11,917
Sign	21,446
GR	15,824
IR	14,381
Release	7,635

After the preprocessing step, auditors perform an in-depth process discovery analysis on the event log dataset for the underlying business process. The objective of this step is to better understand the business process and discover the paths employees are following to perform a transaction. Figure 12 illustrates the detailed process map for the P2P process.

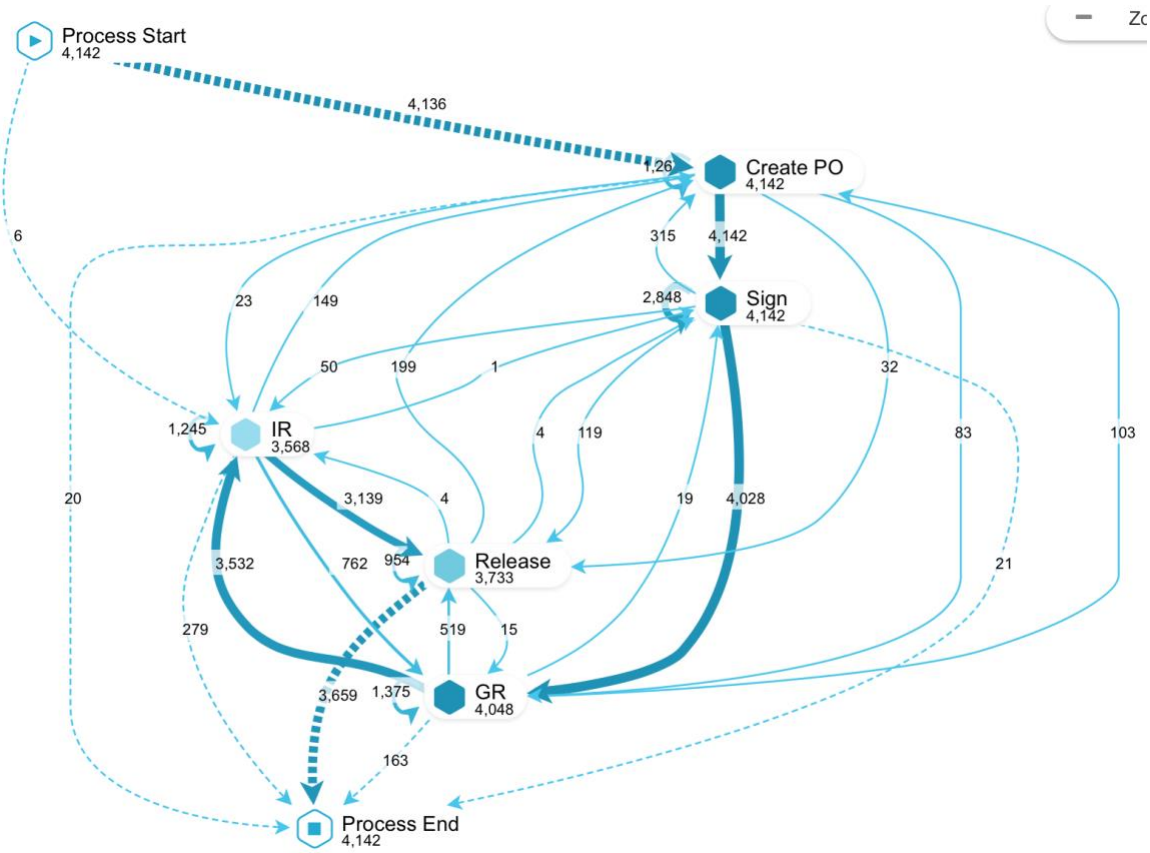


Figure 12. P2P Process Map with 100% of Paths

Even though the majority of process instances follow the standard P2P process, by examining the process map, auditors can see that there are business rules and controls that are being violated. For example, there are 91 Create PO activity that was not followed by a Sign activity, indicating improper authorization. Variant analysis is performed to examine the population and identify risk factors. In this application of the framework, several risk factors based on the process flow were identified and filters were developed to examine them. There were four major categories identified as risk factors: missing key activity, problematic order, weekend activity, and segregation of duty. Additionally, auditors must

give a risk level or weight for each risk factor based on their professional judgment so that the exceptional process instances can be prioritized. Table 8 provides a list of the process-related risk factor categories and the filters developed to examine those risk factors.

Table 8. P2P Process-Related Risk Factors and Filters

Risk Factor Category	Sub-Category	Filter	Risk Level
Missing Key Activity	Missing GR activity	Any process instance missing GR	High
	Missing IR activity	Any process instance missing IR	High
	Missing Release activity	Any process instance missing Release	High
Problematic Order	Unusual start for a PO	Process instances starting with IR	Medium
	Payments before all goods are received	Payment before GR	High
	Ending with Create PO	Process instances ending with Create PO	Low
	Ending with GR	Process instances ending with GR	Low

Segregation of Duty	SOD violation	The employee who created the order also released the order	High
Weekend Activity	Unauthorized weekend activity	Process instances with activities happening on weekends	Medium

The risk level for each risk factor was given based on the feedback from 6 auditors from the Big 4 accounting firms and 2 other major firms. The risk factors were sent to them for their professional risk assessment. The results of the assessment were averaged out for the purpose of this study. The final risk levels are found in Table 8.

Notable Process Instances

The process-related filters were then applied to the entire population of process instances. The result of this application is broken down amongst the four risk factors categories. First, for the missing key activity category, the first filter applied was related to process instances missing a GR activity. The risk associated with this filter is that the organization might be billed goods they did not receive. In this case, there were 94 process instances found that were missing GR. However, every process instance of those 94 did not have an invoice associated with it, and it also had a PO value of $0. This indicates that all the 94 process instances were nullified. Therefore, those process instances were not considered notable even though they violated a business rule. It is worth noting that applying filters carelessly without considering other factors may not be very beneficial to auditors and may actually affect the results they're seeking to achieve.

The second filter applied in the missing key activity category is missing IR activity. The risk is that the organization might be paying for an invoice they did not receive. There were 574 process instances in 130 variants that were found missing an IR activity. However, missing IR alone is not sufficient to be considered a notable process instance if the organization did not receiver goods for the PO. Therefore, the filter was modified to include process instances that were missing an IR activity but at the same time have a GR activity. The modification resulted in a more focused and relevant notable process instances, which got down to 480 notable process instances in 108 variants.

The last missing key activity filter was for process instances missing Release activity. Any PO created needs to be released to the accounts payable department for payment and recorded in the general ledger. If a Release activity is not present, then this might indicate that the PO did not transfer to the accounts payable department and may not have been payed. In the event log dataset, it was found that 409 notable process instances were missing a Release activity. Those process instances were part of 336 variants. Table 9 presents the results of missing key activity filters.

Table 9. Missing Key Activity

Sub-category	Process Instances	Variants
Missing GR Activity	94	22
Missing IR Activity	480	108

Missing Release Activity	409	336

Second, the problematic order category filters were applied. The first filter of this category relates to process instances with unusual starting activity. In this study, there were 6 notable process instances that started with an IR activity instead of the standard Create PO like the rest of the population. This indicates that the organization received those invoices then it created PO to match them, which is a violation of the business rules.

The second filter applied in this category was for process instances with payments before all goods were received. This filter is concerned with the risk that the organization might be paying invoices that have not been fulfilled yet. There were 176 process instances that are part of 155 variants in which matched this filter. It is important to note that this number can be reduced by applying other analytical procedures such as considering if the PO has been fulfilled over the long term by examining the GR and payment totals.

The last two filters of the problematic order category relate to incorrect ending activity to the PO. In this case, there were two filters: process instances ending with Create PO and process instances ending with GR. Both of these filters are not considered high risk level, however, they still constitute an irregular and problematic order from a process perspective, since all process instances should end with a Release activity so that it can be paid. There were 20 notable process instances ending with Create PO and 163 notable process instances ending with GR. Table 10 presents the results of problematic order filters.

Table 10. Problematic Order

Sub-category	Process Instances	Variants
Unusual start for a PO	6	3
Payments before all goods are received	176	155
Ending with Create PO	20	14
Ending with GR	163	142

Third, segregation of duty category filter was applied. This category was mainly concerned with one risk factor. The focus was on examining a segregation of duty violation on whether there were process instances that had the same employee who created the PO also released it. This filter resulted in 36 notable process instances that are part of 33 variants in which segregation of duty was violated. Table 11 presents the results of segregation of duty filter.

Table 11. Segregation of Duty

Sub-category	Process Instances	Variants
SOD violation	36	33

Finally, weekend category filter was applied. This filter examines process instances that have unauthorized activities happening on a weekend. Even if some process instances are following the standard process flow, it still may be flagged if the timestamp of certain activities (i.e. GR and IR) is on a weekend. This could be of concern to auditors if weekend activities are suspicious. In this case, there were 551 notable process instance that contained activities happening on weekends. Table 12 presents the results of weekend activity filters.

Table 12. Weekend Activity

Sub-category	Process Instances	Variants
Unauthorized weekend activity	551	296

As a result of applying all process-related filters to the entire population of process instances, 1,346 unique notable process instances (out of 4,142) are identified. These notable process instances will be the beginning point for the existing analytical procedure element of the framework and will be filtered down to exceptional process instances. The next section will discuss in detail the results of the existing analytical procedures element

of the framework.

3.4.2. EXISTING ANALYTICAL PROCEDURES ELEMENT

The risk factor of the previous stage of the framework relate to process risk factors, and not the other factors that are non-process mining related. This stage is of the framework allows auditors to narrow down the number of notable process instances into the highly problematic exceptional process instances. For that purpose, filters for non-process-related filters were developed. The development of those filters was based on the WCGW for a standard P2P business process and auditors risk assessment. There were also filters that resulted from risk factors identified in the process discovery step. The filters identified in this stage were grouped into two categories: missing values, and 2-way match. As with the case of the process-related filters, the non-process related filters in this stage of the framework were given a risk level or weight for each risk factor based on auditors' professional judgment. This allows for the exceptional process instances found in this stage to be prioritized. The same methodology that was used in the previous stage for assigning risk level for the filters was followed in this stage as well. Table 13 provides a list of the other process risk factor categories and filters.

Table 13. P2P Other Process Risk Factors and Filters

Risk Factor Category	Sub-Category	Filter	Risk Level
Missing Values	Missing PO value	Any process instance missing PO value or equals $0	High
	Missing PO quantity	Any process instance missing quantity value for PO	High
	Missing Invoice	Any process instance missing invoice value	High
2-Way Match Violation	Unmatched PO and goods values	Any process instance with PO and GR values that do not match	High
	Unmatched PO and goods quantities	Any process instance with PO and GR quantities that do not	High

		match	
	Unmatched goods and invoice values	Any process instance with GR and IR values that do not match	High

Since this analysis is done on already notable process instances, the risk factors of this stage are all at a level of high-risk. Additionally, the filters are guided by the audit objective, auditors risk assessment, and the findings of the previous stage of the framework. This means that if auditors find a relatively large number of notable process instances, the filters used in this stage might be different or larger in number. This also depends on the type of violations of the notable process instances.

Exceptional Process Instances

The main objective of this stage is to determine the outliers of the notable process instances that are more likely to be of high risk and problematic. Therefore, the filters of this stage were applied to the entire set of notable process instances. The first category of risk factors was to identify any process instance with missing values. The first filter applied was for process instances missing PO values. These process instances run the risk that either the employee did not declare the value of the PO or it was not recorded in the information system. From the notable process instances, there were 252 that this filter applied to. It should be noted that in this stage, the number of variants is not relevant since the focus is on process instances and the tool used is not a process mining tool. This is important because it demonstrates how event logs with added relevant information

contributes to the effectiveness of the audit and allows for additional analysis that would be difficult or not possible without.

The second filter applied in the missing value category is for POs with a quantity of 0 while the it had a monetary amount greater than $0. This filter addresses similar risks as the previous filter. Even though the previous filter can act as a "compensating filter", the notable process instances should be examined for other business rules violations. There were only 2 POs missing values for quantities.

The last filter in the missing value category is for process instances missing an invoice value. This filter is concerned with the risk associated with a PO having an invoice value of $0, while its PO and GR values are of greater amounts. This might indicate errors in recoding invoices or payments maid. In this study, there were 18 notable process instances that this filter applied to. Table 14 presents the results of missing key activity filters.

Table 14. Missing Values

Sub-category	Process Instances
Missing PO value	252
Missing PO quantity	2
Missing Invoice	18

Filters for the other risk factor category were applied. These filters related to 2-way match risk. In some instances, the 2-way match business rule was not adherent to. This is a high-risk category since having a 2-way match control ensures that errors and misappropriation of assets are mitigated. First, process instances with unmatched PO and GR values are filtered out. This resulted in 622 process instances. It should be noted that process mining alone cannot determine if repetition in a GR activity is indicative of goods being received over installments or is a mismatch between the goods being received and its associated PO without examining their monetary values. Therefore, additional value analysis is required, such as the one applied in this filter, to determine that.

Second, the PO and GR quantities were tested for all notable process instances to ensure that all goods were received from suppliers and no pending goods for completed POs. This filter resulted in 652 process instances. The same comment for the previous filter applies to this filter as well.

Finally, the last filter applied relates to unmatched values of invoices received from suppliers with values of goods received. The risk associated with this filter is overpayment for POs. Note that the value of goods and invoice might not always match due to additional shipment and other costs unaccounted for in the original PO, and therefore should be tested for violation of an accepted variance, such as 5% over PO value. There were 353 process instances that violated this filter. Table 15 presents the results of missing key activity filters.

Table 15. 2-Way Match Violation

Sub-category	Process Instances
Unmatched PO and goods values	622
Unmatched PO and goods quantities	652
Unmatched goods and invoice values	353

The application of the existing analytical procedures resulted in a total of 814 unique exceptional process instances. Auditors might still find this number to be large and therefore will consider focusing on exceptional process instances with material amounts. Therefore, applying a threshold allows auditors to focus on investigating process instances that could have a potential impact on the financial statements. This study applied a threshold of $5,000 as a filtering method. The final results for exceptional process instances with a materiality threshold is 457 process instances. This shows that by following the framework, it was possible to narrow down the number of exceptional process instances to 457 out of 4,142, which is about 11% of the entire population of process instances. The final stage of the framework is to rank the exceptional process instances based on a prioritization method. The next section will discuss the prioritization element of the framework.

3.4.3. PRIORITIZATION ELEMENT

The final part of the framework is the prioritization of exceptional process instances. The prioritization method used is based on the risk score calculated for each

exceptional process instance. The risk score is calculated based on three factors: the number of filters violated in both the process mining stage and the other analytical procedure stage, the weight given to each filter, and the dollar amount of the PO. These three factors are chosen to calculate the risk score because they consider the three most important elements that auditors look for when analyzing audit results (frequency of violation, importance of violated controls, and materiality impact). In this study, the dollar amount was normalized so that the risk score would not be in the millions ($1,000,000 = 1, $500,000 = 0.5, etc.) and therefore, easier to recognize. Table 16 provides a list of the top 10 exceptional process instances with the highest risk score.

Table 16. Highest Risk Score Process Instances

Process Inst.	Missing IR	Missing Release	Unusual start	Pay before GR	Create PO ending	GR ending	SOD violation	Weekend	Missing PO value	Missing PO quantity	Missing invoice	Unmatched PO & GR	Unmatched quantities	Unmatched GR & IR	Violation Score	Monetary Value	Risk Score
	3	**3**	**2**	**3**	**1**	**1**	**3**	**2**	**3**	**3**	**3**	**3**	**3**	**3**			
88702		X						X				X	X		11	$11,579,094	**127**
88814		X						X				X	X	X	14	$7,882,137	**110**
87646								X				X	X	X	11	$3,137,867	**35**
87639								X				X	X	X	11	$2,659,998	**29**
89106								X				X	X		8	$3,120,400	**25**

89465						X						X	X	X	10	$1,507,415	**15**
88749						X						X	X	X	10	$1,179,759	**12**
89503	X	X		X								X	X		13	$877,638	**11**
87640								X	X			X	X	X	14	$729,189	**10**
88816		X										X	X	X	12	$686,957	**8**

The number beneath each filter in Table 16 is the weight given to that filter based on auditor's professional judgement. The exceptional process instance (88702) that is ranked the highest in terms of risk score (127) did not actually have the highest violation score (11). Process instance (88814) had the highest violation score (14) which is calculated based on the number of filters that applied to it. This is due to considering the monetary value of the process instance when ranking them, since 88702 had a PO value of $11,579,094 while 88814 had a PO value of $7,882,137.

The process instances risk prioritization framework provides auditors the riskiest transactions. However, auditors still need to use their professional judgment to determine how many of the prioritized process instances they need to examine thoroughly and perform a substantive test of details on. For example, auditors may determine that they need to perform substantive test of details on 50 process instances with the highest risk score.

Chiu et al. (2018) discussed four different prioritization methods for ranking high-risk transactions. Each of the four prioritization methods found in that study emphasizes

different aspects of the risk associated with the transaction. For example, based on their professional judgement, auditors can either focus on the risk score of transactions without considering the monetary value, transactional value without considering the number of violations, or a combination of both. However, this study implemented a prioritization method using a suspicion function formula that considers the number of business rules violations, the weight of each given business rule, and the monetary amount of each process instance. This prioritization method is more objective in considering different risks associated with the exceptional process instances.

3.4.4. FRAMEWORK COMPARISON

The process instances risk prioritization framework was demonstrated on a real-life full population of process instances found in the procurement process of a not-for-profit national organization. The full population was comprised of 4,142 process instances. After applying the first part of the framework, which is the process mining part, it highlighted 1,346 notable process instances. However, applying the filters of the second part of the framework, the existing analytical procedures, resulted in narrowing down the results of problematic process instances to 814 exceptional process instances. After considering the material impact of these exceptional process instances, it was suggested to use a threshold of $5,000. This resulted in focusing only on highly problematic process instances that have a material amount, which were 457 (about 11% of the entire population). Therefore, by using this framework, auditors can be guided to the riskiest and highly problematic process instances as opposed to choosing from a random sample in a traditional audit.

On the other hand, Chiu et al. (2018)'s study demonstrated the integration of process mining into the auditor's risk assessment process by combining process mining results with a corresponding transaction value of each process instance. The focus of that study was purely on process mining and not on other existing analytical procedures. Therefore, the process instances risk prioritization framework expands on Chiu et al. (2018)'s framework by adding a second stage of filtering that that study lacked. The result of Chiu et al. (2018) study was 3,918 notable process instances out of a population of 9,187. That study also applied a threshold of $5,000 which filtered the results down to 1,227. Table 17 provides a comparison of the results of this study and Chiu et al. (2018) study.

Table 17. Results Comparison with Chiu et al. (2018)

	Process Instances Risk Prioritization Framework		**Chiu et al. (2018)**	
	Count	*Percentage*	*Count*	*Percentage*
Notable Process Instances	1,346	32.5%	3,918	42.6%
Exceptional Process Instances	814	19.7%	--	--
Threshold	457	11.0%	1,227	13.4%

Having a second stage of filtering that complements process mining concentrates auditors' focus on the process instances that are most likely to be problematic. This is a more objective way of guiding auditors to the riskiest process instances.

3.5. CONCLUSION

This paper introduces a methodology that provides auditors with guidance to objectively identify and prioritize the riskiest process instances. The process instances risk prioritization framework is based on applying process mining techniques on an event log extracted from the organization's information system to detect anomalies. These anomalies are then filtered using other analytical procedures to identify high-risk exceptional process instances. The exceptional process instances are then prioritized based on a calculated risk score. This combination of process mining with other analytical procedures is unique to this study.

The aim and contribution of this study is to provide auditors with guidance on the use of process mining in conjunction with existing analytical procedures to identify exceptional transactions that would require further investigation. This solution allows auditors to focus on process instances that are likely to be considered high-risk, reduce the risk of failing to detect material misstatement, and enhance audit effectiveness. Furthermore, the identification and prioritization of such risky process instances help with the information overload problem that entails population testing.

The identification and prioritization of exceptional process instances depend on the filters developed in the different stages of the framework and the weight given to each filter. A limitation of this study is that the results could differ depending on the filters developed and the weights given.

Even though this study supports the application of process mining prior to other

existing analytical procedures, future studies could examine whether the prioritized exceptional process instances would differ if process mining is applied in the final stages of the framework as opposed to the early stages and compare the results. Furthermore, future work can include the implementation of the risk score methodology found in this study to a continuous process mining solution as a way to allow or block transactions based on their risk score.

CHAPTER 4: CONTINUOUS PROCESS MONITORING

4.1. INTRODUCTION

In the wake of several highly-publicized corporate scandals, the Sarbanes-Oxley Act of 2002 (SOX) was enacted by the United States Congress, and consequently, the focus on internal controls has tremendously increased. SOX Section 404 requires management to assess the effectiveness of their internal control system, and for auditors to attest to management's assessment. Since its introduction, the focus on fraud and fraud detection/prevention was put to the forefront and the improvement of internal controls were of major importance to reduce any risks of fraudulent transactions.

The importance of having an adequate internal control system cannot be overstated. Prior research has found that when internal controls are weak, there is an increased likelihood of earnings manipulation by management (Chan et al. 2008; Ashbaugh-Skaife et al. 2008). Moreover, effective internal control system can help companies achieve their financial goals, prevent loss of resources, keep accurate recording of transactions, and comply with laws and regulations by preparing reliable financial statements (Ernst & Young 2002). Hence, maintaining an effective internal control system is regarded as highly important to management.

With the passing of SOX and the digitization of the economy, internal control evaluation has changed dramatically from being mainly used by management to endure operational efficiency, to being a legislative requirement. Management has modified its efforts and focus to comply with SOX by emphasizing the importance of assessing, developing, and maintaining an effective and efficient internal control system

(Rikhardsson and Kræmmergaard 2006). Post-SOX studies have emphasized the need for formal control assessment and compliance methods utilizing computer aided tools to comply with rules and regulations and improve the effectiveness of the internal control system.

Traditionally, the testing of controls has been performed on a retrospective and cyclical basis, often many months after business activities have occurred. The testing procedures have often been based on a sampling approach and included activities such as reviews of policies, procedures, approvals, and reconciliations. With today's real-time economy and the advancements in technology, it is recognized that this approach only offers auditors a narrow scope of evaluation and is often too late to be of real value to business performance or regulatory compliance. Therefore, the motivation behind this study is to reduce the time delay that traditionally manifests its self between the occurrence of a business related event and its analysis. This can be achieved by applying continuous monitoring methods to business processes. The reason for that is because continuously monitoring business processes increases the information value by investigating events simultaneously or shortly after their occurrence (ISACA Standard Board 2002). In other words, the availability of real-time data allows for exceptional cases to be identified and dealt with before they lead to issues (Selig 2017). Additionally, having a preventative focus is fundamental to achieve sustainable compliance (Agrawal et al. 2006). Today's business environment allows for the adoption of continuous analytical monitoring-based assurance, which is an outcome of the fundamental transformation in business operations and control that stems from the electronization of firms through the widespread use of ERP systems (Vasarhelyi et al. 2004).

Many process mining studies in the past have focused on implementing process mining as a discovery tool for auditing and for detecting anomalies or deviations in business processes (i.e. Van der Aalst & Medeiros, 2005; Van der Aalst et al. 2010, 2011; Jans et al. 2011, 2013, 2014). However, the use of process mining on a continuous basis to monitor business processes and provide assurance, to our knowledge, hasn't been found in the auditing literature. This paper's contribution to the auditing literature is by developing a novel approach for monitoring assurance that combines the advantages of continuous monitoring with those of process mining. Auditors can actively detect and investigate deviations and exceptions as they occur along the transaction process by continuously monitoring business process controls and testing transactions, rather than react after the exceptions have long occurred. Any transaction that violates a set of business rules would be intercepted or flagged by the system until investigated by an auditor. This continuous monitoring using rule-based process mining approach provides a high level of assurance about the operating effectiveness of controls throughout a business process. Basically, this study is attempting to answer the research question of how can the time delay between the occurrence of a business operation related event and its analysis be reduced. Additionally, can process mining be implemented automatically or does it always have to be manual?

The paper will be organized as follows: section 2 will provide a background and literature review on continuous auditing and assurance, control monitoring and compliance verification, and process mining. This will be followed by section 3, which describe the methodology and general framework developed in this study. Section 4 will include the data used to demonstrate the methodology, the analysis, and some application scenarios. Finally, section 5 will be the conclusion for this paper.

4.2. BACKGROUND

The information system of many corporations, especially large ones, can be a blend of legacy systems, middleware, and different ERPs. This complex integration and real time nature of the transactional data in many firms increases the likelihood of discrepancies to occur, whether it be errors or fraud. This is posing significant challenges to managers and auditors to re-engineer business processes and adopt new technologies to develop methods and tools to continuously monitor and improve internal controls. Fortunately, information systems encompass several control features that help prevent several forms of errors. However, in many cases, these controls are ineffective due to several reasons. First, not all system controls are switched on by an organization, some firms choose to keep certain controls switched off or deactivated to allow for flexibility in conducting business operations. Second, continuous monitoring and control is absent. Having a system in place that continuously monitors the effectiveness of internal controls could provide firms with more reliable data to safeguard their assets. Finally, some authorized or unauthorized users may have the authority and/or ability to bypass or override certain controls (Islam et al. 2010).

Also, since the responsibility for adopting sound accounting policies, maintaining an adequate internal control system, and making fair representations in the financial statements is on management, it is only logical to propose a system that aids them in this task. A model for continuously monitoring the effectiveness of internal controls would facilitate the transfer of internal control knowledge to a manager, thereby supporting their decisions from an internal control perspective (Arens et al. 2000, Changchit et al. 2001).

In addition, having effective and reliable systems that aid managers in understanding and monitoring their internal controls are feasible. Such systems save the firm time and money by detecting weaknesses in internal controls rapidly and maintain an effective internal control system.

4.2.1. CONTINUOUS AUDITING AND CONTINUOUS ASSURANCE

Interest in continuous auditing has been progressively increasing amongst practitioners and academic researchers ever since its introduction by Groomer and Murthy (1989) and Vasarhelyi and Halper (1991). This increased interest suggests the need to develop improved auditing methodologies that take advantage of new technologies, especially in the real time economy (Chan and Vasarhelyi 2011). Many studies have shed light on the technical aspect of continuous auditing (Kogan et al. 1999; Woodroof and Searcy 2001; Rezaee et al. 2002; Vasarhelyi et al. 2004), and others discussed the feasibility of implementing it in organizations and its impact on audit practice (Alles et al. 2002; Alles et al. 2004; Elliott 2002; Vasarhelyi and Halper 2002).

What differentiates continuous auditing from traditional auditing is its changes to three key aspects: nature, timing, and extent (Vasarhelyi and Halper 1991). continuous auditing changes the "*nature* of the audit" as internal control monitoring and transaction data testing are used on a continuous bases to evaluate management's assertions instead of performing manual internal control and substantive detailed testing periodically. This continuous evaluation of controls and processes are the cornerstones of their study. The second change of continuous auditing to traditional auditing is in regards to the "*timing* of the audit". In a continuous auditing environment, internal controls monitoring and transaction data testing occur simultaneously, which is necessary to support real time

assurance (Rezaee et al., 2001). This is in contrast to traditional auditing where internal control testing occurs in the planning stage while substantive detail testing occurs in the fieldwork stage of the audit. This paper's methodology combines both continuous control monitoring and continuous data assurance to provide real time assurance on the effectiveness of controls for the specified business process. The third change to traditional auditing is to the "*extent* of the audit". Continuous auditing allows auditors to automate the testing the whole population of data rather than having to rely on manually testing a sample. Auditors traditionally have to rely on sampling techniques when testing internal controls and transactional data due to manual testing's labor and time intensiveness. However, the consideration of the whole population of transactions in testing can enhance the effectiveness of an audit and increases the probability that material errors, omissions, fraud, and internal control violations may be detected (Chan and Vasarhelyi 2011). These three key differences of continuous auditing with traditional auditing are the main drivers for adopting continuous auditing and monitoring methodologies with process mining to provide real time assurance for the effectiveness and compliance of business process controls.

In addition to the advantages that continuous auditing has over traditional auditing when it comes to the nature, timing, and extent of an audit, continuous auditing can provide the opportunity for an audit to proactive rather than reactive. This means that instead waiting till the end of the audit period to audit accounting information and allow for material errors, omissions, or fraud to go undetected for months, continuous auditing involves the implementation of continuous control monitoring, continuous risk monitoring and assessment, and continuous data assurance that allow auditors to actively detect and

investigate violations as they occur rather than to react after the violation has long occurred (Chan and Vasarhelyi 2011). Hence, violations can be intercepted and blocked before the completion of a transaction until investigated by an auditor to prevent errors or fraud.

Along with implementing continuous auditing in organizations, accounting researchers have urged towards "continuous assurance", which consists of continuous auditing and continuous monitoring. Alles et al. (2003) defines it as "technology-enabled auditing which produces audit results simultaneously with, or a short period of time after, the occurrence of relevant events". Hence, continuous assurance provides decision makers with assurance over a continuous stream of data. This relies on capturing information that relate to transactions and processes, which are continuously monitored to identify any discrepancy between actual and expected results. The methodology proposed in this study which involves continuous monitoring using rule-based process mining techniques is a demonstration of continuous assurance for a specific business process. This methodology is stemming from the need to seek new audit evidence that auditors can utilize to improve audit quality. Process mining is a new type of audit evidence that can be a great benefit to auditors. However, new methods are needed to analyze the evidence gathered by process mining.

4.2.2. ABSTRACTED LAYER IMPLEMENTATION FOR CONTROL MONITORING AND COMPLIANCE VERIFICATION

The implementation of information systems with only digital transactions has created opportunities for auditors and researchers to take advantage of electronic evidence that was not available previously, to perform effective and efficient audits (AICPA 1997; Williamson 1997; Lavigne 2003) Therefore, many studies have proposed new frameworks

that would be appropriate for IT centric information systems or proposed new methodologies for compliance with laws and regulations. For example, Namiri and Stojanovic (2007) proposed an approach for modeling and implementing internal controls in business processes by introducing an abstraction layer above a business process. They focus on application controls, which are controls that relate to each computer-based application and are specific to that application. The objectives of such controls are to ensure the completeness and accuracy of the records and the validity of the entries made within the application. The advantages of introducing an abstraction layer above a business process include four points (Namiri and Stojanovic 2007):

- Formal methods can be used for the verification of a business process's compliance.
- This formalization enables the compliance to be performed automatically based on the current state of a process.
- Changes made to the controls will not affect the design and execution of the original business processes.
- Allows non-experts to build on top of the domain model provided to design controls for business processes.

By implementing an abstraction layer, each application control has at least one recovery action designed for it, which reacts on the violation of a control. The recovery action does not change the designed business process logic, rather it blocks the transaction and send a notification to an assigned responsible agent. This semantic based approach for internal control compliance proposed in that study would be a valuable building block for demonstrating how internal control effectiveness can be assessed in a formalized way.

Another study that proposed a different methodology for internal control compliance was by Borthick (2012) where the author illustrated how the stages of continuous auditing proposed by Chan and Vasarhelyi (2011) might be implemented in a highly automated procure-to-pay process using the Krishnan et al. (2005) notation for representing controls in business process diagrams. The study was based on the idea that for auditors to provide real-time assurance, they must rely on continuous auditing. An interesting point about this study is that it showcased how continuous auditing can be used to monitor and assess internal control compliance.

Continuous monitoring of business process controls has also been demonstrated in a pilot study by Alles et al. (2006). The study showed how an independent monitoring system running on top of the enterprise information system could perform audit tasks on a continuous basis. This study focused on configurable application controls and tested them by retrieving the control settings stored in the organizational information system and verify that they match a well-defined benchmark. This approach is very beneficial since it can be accomplished by just having read-only access to the organizational information system, which provides a very strong evidence since it actually confirms that the control is indeed what it has to be.

The common theme amongst the studies discussed above is the introduction of a layer on top of the business process to assess and monitor internal control compliance. Having such layer provides many benefits, such as providing formal assessment of controls and changes made to the controls will not affect the design of the underlying business process. Therefore, the proposed method in this paper will rely on implementing an abstracted layer

on top of a business process for testing internal control effectiveness and compliance verification.

4.2.3. PROCESS MINING AS AN APPROACH FOR MONITORING INTERNAL CONTROL COMPLIANCE

Regulations require management to implement an effective internal control system in their organization. Assessing controls involves checking, comparing, monitoring, and taking action when deviations from the modeled design are found (Rikhardsson and Kræmmergaard 2006). As a way to deal with the complexities of today's information technology environment, process mining was introduced to validate information about companies and their business processes.

Process mining is a tool originally developed by computer scientists to aid in identifying and analyzing business processes (Jans et al., 2013). It achieves this by providing techniques and tools for discovering process, control, data, organizational, and social structures from event logs (Van der Aalst & Medeiros, 2005).

The event log is the starting point for any process mining analysis. The log contains events, and each event refers to an activity that can be viewed as a well defined step in a process. Additionally, each activity relates to a particular case or a process instance. Therefore, a case contains a sequence of events that is unique to that case. Moreover, other information can be stored in the event log, such as the initiator of the activity or the resource, the timestamp of an event, or financial value of a certain event in a transaction sequence.

Process mining of event logs is a method for understanding the complex operation of business processes. The data analyzed by process mining consists of the event log that

the auditor constructs from records maintained by a business's information systems. The event log includes information about the activity, the identity of the person who performed the activity, the time of execution, and other contextual information. Process mining is a unique audit tool because it focuses on the path of transactions and not directly on the validation of the values in the associated process. This makes it a powerful tool for tests of controls, such as those for segregation of duties (Jans et al. 2014). Not only that, but process mining can be applied to the whole population of data instead of a sample as in traditional auditing procedures, which increases the reliability of the results since auditors are testing all instances.

However, along with auditing the path of the transaction, auditors need to consider the values underlying the transactions audited. Therefore, to have a comprehensive internal control compliance framework, it needs to include both aspects. These two aspects will be discussed in this study's framework.

4.2.4. COMPLIANCE VERIFICATION PROCESS MINING TECHNIQUES

There are three broad process mining techniques that deal with compliance verification: process discovery and visualization, conformance checking and delta analysis, and logic-based property verification (Caron et al. 2013). The first technique, process discovery and visualization, is concerned with understanding the real business process and discovering how each step throughout a business process is being executed. It also includes visualizing the different aspects of a business process, such as internal controls or social networks, to detect violations and anomalies. The second technique, which is conformance checking and delta analysis, focuses on comparing the actual process flow with a verified model of the business process to detect any inconsistencies and internal control violations,

for example. The third technique, logic-based property verification, which is the main focus of this study, relates to analyzing specific process properties included in the business process. For example, analyzing the timing of a certain activity prior to another activity, and other activity preconditions. The first two broad techniques provide a general holistic view of the business process, while the third technique is used to provide analysis of specific process instances.

4.3. METHODOLOGY

The architectural methodology for implementing continuous monitoring with process mining techniques is based on implementing an abstracted layer on top of the business process that would continuously monitor the activities throughout a transaction and prevent or flag any violations (Vasarhelyi et al. 2004). The event stream of the information system is used as an input for the monitoring layer, which consists of an adapted rule-based process mining technique. So, instead of relying on "after the fact" process mining techniques, the system would flag any transaction that does not conform with the approved model for that business process. Hence, logs can be automatically generated and process instances are automatically mined in real-time for deviations and assuring compliance. In addition, the continuous monitoring layer can support the implementation of not only preventative controls, but also timely corrective controls that correct deviations when they are detected.

Prior to discussing the framework, there needs to be an understanding that to implement such monitoring layer for continuous assurance, it requires two key components: an IT structure that facilitates data gathering, and an analytic monitoring methodology to support continuous monitoring (Vasarhelyi et al. 2004). These two key

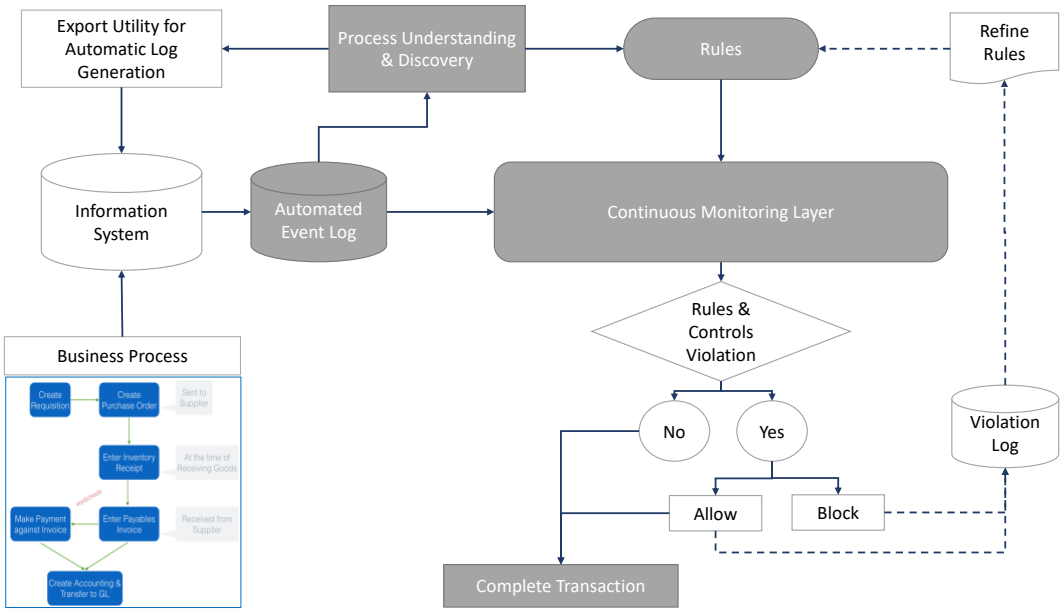
components are critical for implementing continuous assurance as suggested by Chan and Vasarhelyi (2011).

4.3.1. FRAMEWORK

This section provides a conceptual illustration of the continuous monitoring layer using rule-based process mining techniques, along with examples of rules violated for transactions that are part of a P2P process.

Figure 13 illustrates the implementation of a continuous monitoring layer on top of a business process that utilizes rule-based process mining techniques to provide continuous monitoring and assurance on the effectiveness of controls. This framework and its different components will be discussed in the following subsections.

Figure 13. Continuous Monitoring of Business Controls Using Rule-Based Process Mining Technique

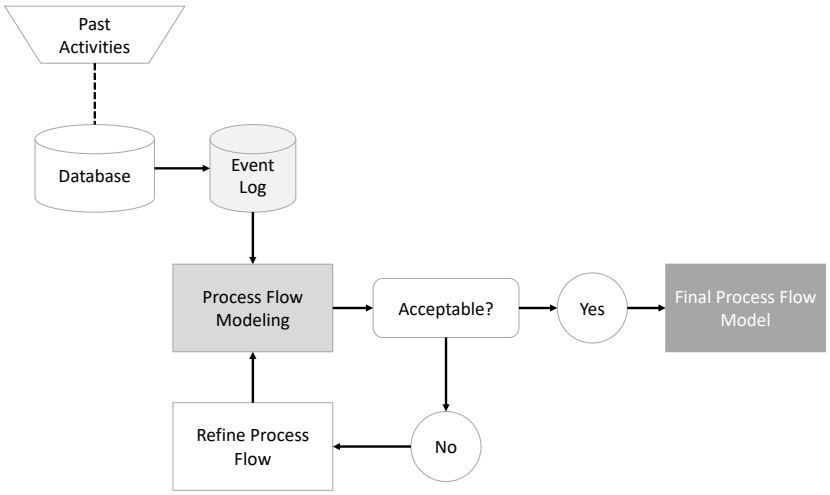


Process Understanding and Discovery

Prior to the implantation of a continuous monitoring layer, there needs to be a clear understanding of how transactions of the underlying business process are being executed. In a traditional audit setting, auditors gain understanding of a business process through conducting interviews and doing walkthroughs. Even though this provides auditors with a certain degree of understanding, it is only limited to what management and employees declare to them during the interviews, and to what transactions they observe during their walkthroughs. Hence, a better and more thorough way to gain an understanding of the underlying business process is to allow the transactional data itself to inform the auditor about all the process methods these transactions are being executed. This can be achieved by using process mining as a tool to conduct process discovery analysis.

Internal control processes are a set of process flows, each containing required control activities. The first aspect of the framework, process discovery, starts with examining all past transactions of a current business process to establish the path for that process. This is done by modeling past transactions using logs (a bottom-up approach) that would then be used as a baseline to model required process flows for routine company transactions. Figure 14 depicts the steps of process discovery and flow modeling.

Figure 14. Process Discovery and Flow Modeling of a Business Process



The information system of a company records certain transactional activities in logs that are stored in different tables. These activities include controls over initiating, authorizing, recording, processing, and reporting significant accounts, disclosures and related assertions in the financial statements. The PCAOB requires auditors to understand the flow of transactions related to relevant assertions as the ones just mentioned (PCAOB Auditing Standard No. 5). The logs used for process discovery should include information about the activity, the identity of the person who performed the activity, the time of execution, and other contextual information (Agrawal et al. 2006). However, in order to create an event log, auditors need to understand what are the activities that make up the examined process and what case will be followed throughout the process to create a complete transaction (Jans et al. 2011). In addition, if the information system is not process aware, meaning that event logs can be created and extracted automatically, then a manual process needs to be undertaken to extract the necessary information from multiple tables

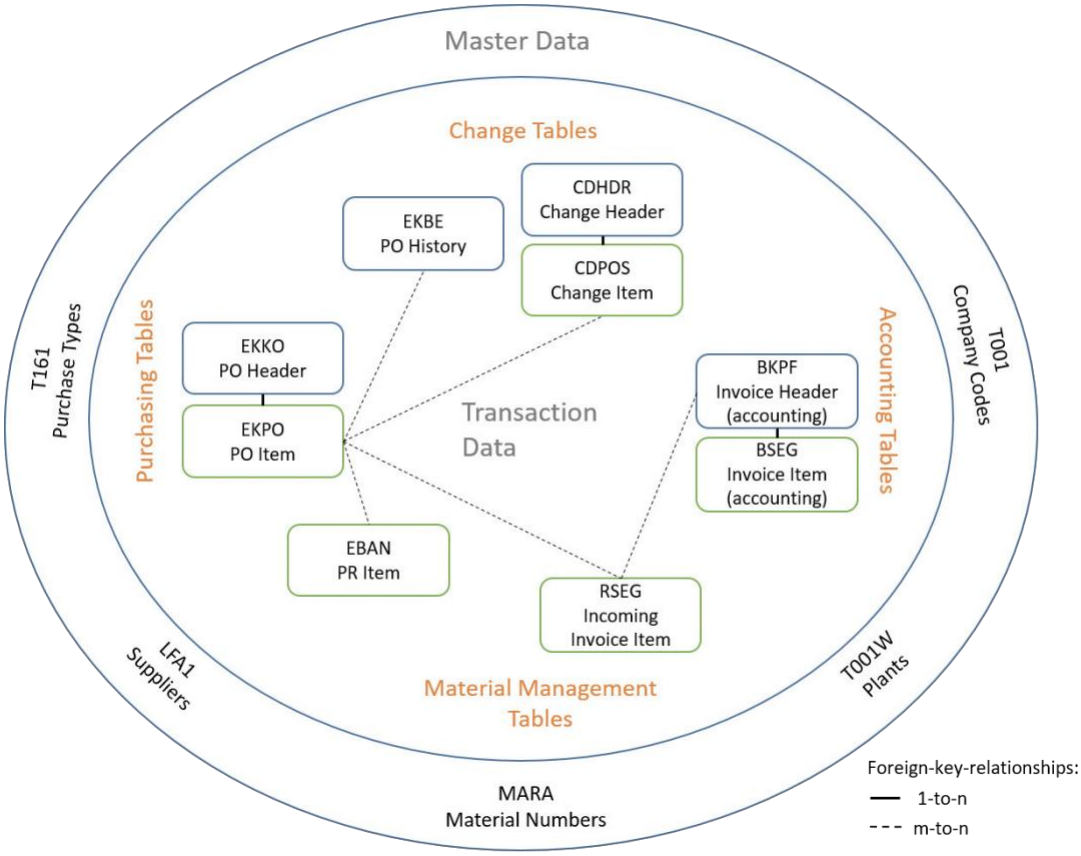
that constitute the underlying process. For example, more than 30 tables can be used to extract data from and construct an event log in an SAP ERP system for the procure to pay process (P2P). Table 18 provides an example of some of the tables used for P2P event log generation.

Table 18. Example of Tables Used for P2P Log Generation

SAP Table	Technical Code	Notes
Purchasing Document Item	EKPO	Used as the case table
Purchasing Document Header	EKKO	Used to contextualize the process analysis.
History per Purchase Document	EKBE	
Purchase Requisition	EBAN	
Change Document Header	CDHDR	
Incoming Invoice	RSEG	
Accounting Document Header	BKPF	
Vendor Master	LFA1	

It should be noted that knowing the required tables and relations is not intuitive. Constructing an event log is a trial and error process until the correct tables and attributes are identified. Figure 15 provides an overview of the most relevant tables in an SAP system and their foreign key relationships for the P2P process as depicted by Selig (2017).

Figure 15. Example of Relevant Tables and Foreign Key Relationships for P2P Process



Once the logs are generated, they are analyzed to model and construct the path of past transactions. These process flows will be used as a baseline for modeling the required process flow of transactions for the associated business process. Management and auditors can refine the process flows to ensure that they include adequate internal controls and highlight anomalies or violations that can be of concern to management and auditors. An example of such process model can be found in figure 16 and shows an overview of the steps required to complete a routine transaction in the procure to pay (P2P) process. The most followed path in the business process being analyzed is typically the ideal path for it,

since it is assumed that companies will not allow the vast majority of their transactions to be executed without following the business rules.

Figure 16. Overview of P2P Process

This step allows auditors to establish the "design" of the internal control system in the company and assess its effectiveness from a structure level. This "preprocessing" step is critical, and without it, the implantation of a continuous monitoring layer cannot be effectively achieved.

Automatic Log Generation

Each process consists of multiple steps to complete a transaction. If we take the P2P process for example, it starts with the creation of a purchase order or requisition. The purchase order would then need to be signed and authorized before it is sent to the supplier. Once it's sent, a goods receipt activity and an invoice receipt activity need to be recorded in the information system prior to the purchase order's release and any payment made against it. These steps are shown in figure 4. If any of the critical steps is missing, then that would indicate a violation of internal controls and a possible indication of fraudulent activity.

The steps followed for the P2P process are logged in the information system, and an auditor can apply process mining techniques to detect violations and variants that may be

acceptable. Therefore, the first part to continuously monitoring and assuring the effectiveness of controls using the advantages of process mining is to automatically generate the event log to be used as an input for the continuous monitoring layer. The event log should contain information about the activity, the identity of the person who performed the activity, the time of execution, and other contextual information. This information is stored in different tables in the information system that need to be compiled and aggregated first to create the log. The tables are already determined based on the prior "preprocessing" step of process discovery, which allowed auditors to define the data needed to construct the most suitable event log. Table 1 provided an example of the tables needed to extract the data from to generate the event log. Therefore, to insure a continuous flow of data into the continuous monitoring layer, an export utility implemented as a plug-in on top of the system that would extract the required information from the already determined tables in the information system as the activities are being performed for a transaction and combine them to automatically and continuously generate the log. This export utility is a plug-in that is engineered for that specific information system. Continuous data collection is a necessity for continuous auditing and without a continuous feed of data, continuous auditing cannot be achieved (Vasarhelyi et al. 2004).

However, to be able to capture the data required to automatically generate a thorough event log that covers the whole P2P process, internal controls and business process steps need to be automated and formalized (Alles et al. 2006). Modern IT has provided firms with the opportunity of utilizing converging computer and networking tools to capture business processes information at its source and in the unfiltered and disaggregated form, which makes it possible to measure and monitor business processes at the unprecedented

level of detail in real-time basis (Alles et al. 2006). Traditionally, firms depended on paper-based technology to measure and monitor their business processes. Even though some business process controls could only be accomplished by a human, such as interviewing the client about their reconciliation procedures, research has indicated that more controls involve well-scripted interactions with the client's enterprise system and could be formalized and automated than commonly believed (Alles et al. 2006). For the non-automated controls, information technology can support and facilitate their performance as well as provide a structure for their development and assessment (Rikhardsson and Kræmmergaard 2006). Additionally, if a business process step happens outside the information system, a "step interceptor" can be implemented to extend the existing middleware to intercept it and record it in the log. This is important since process mining, can only analyze data that is captured in the information system. Therefore, formalizing controls or adopting technologies that can record manual controls and activities can make the analysis of processes more thorough and effective since it can capture more activities and data points along the process. However, for the scope of this study, some data point in the log dataset will be simulated as suggested by Caron et al. (2013a; 2013b).

Relevant Rules Identification

The next component is to feed the data obtained from the automatically generated log to the continuous monitoring layer that uses a rule-based process mining technique. Rules have to be defined to compare each activity throughout a transaction against it.

Firstly, there needs to be a clearly defined set of rules that underlie the P2P process so that each step of the process can be measured against. Since business processes contain

dynamic properties, Caron et al. (2013) recommends the use of linear temporal logic when interpreting business rule patterns to develop the rule-based process mining tool. For example, a combination of rule patterns of different types: Activity *a1* should be followed by activity *a2*, and activity *a2* should be preceded by activity *a3*. Other rules might be more static and require only first order logic. For example, a prohibited role-based allocation rule: Originator of role *r1* must not perform activity *a1*. Or, an absolute time rule: Activity *a1* must be performed before a certain time (*T0)*. Hence, it might be sufficient that the event log for some rules only contain timestamps, event and activity identifiers (Caron et al. 2013). But for other complex rules, it might require additional information such the role of the originator or the value of the transaction.

These rules that are defined in this study are adapted from the P2P process' "What Could Go Wrong" (WCGW) and controls provided by one of the Big 4. This set of rules cover many different risk scenarios and levels that a company might face in their P2P process. Additionally, the business rules that are associated with the organization that the event log was extracted from are included. The set of rules include time related events, such as delayed payment of invoice or past the due date. It also considers the roles of the event originator. Table 19 provides a list of rule patterns that are adapted from Caron et al. (2013) for the P2P process that will be used as the base for the rule-based process mining technique:

Table 19. P2P Process Rules

Rule Pattern	Example from the P2P process
An activity of type *a1* must be performed at least once	A *sign* activity must be performed at least once
If an activity of type *a1* is performed then an activity of type *a2* must be performed	If a *goods receipt* activity is performed then an *invoice receipt* activity must be performed
An activity of type *a1* must be started/ completed before/ after/on *t* time units	A *Sign* activity must be started before date of *goods receipt*
A person must not perform both activities for role r1 and activities for role r2	A person must not *sign* and *release* the same *purchase order*
A person must not perform both an activity of type a_1 and an activity of type a_2	A person must not *sign* and perform a *good receipt* activity for the same *purchase order*
A person must not perform all activities of the activity type set s_A	A person must not perform all activities of the P2P process
An activity of type a_1 must be performed under μ	A *good receipt* activity must be performed during regular business hours
A person must perform an activity of type a_1 before/at/after time T (with T referring to time/activity/event)	A person must perform a *release* activity after time T = timestamp of *goods receipt* event
An activity of type a_1 must be performed by a member of role r_1	A *release* activity must be performed by a member of *senior staff*

An activity of type *a1* must not be performed by a member of role *r1*	A *good receipt* activity must not be performed by a member of *senior staff*
The value for event data type *a1* must be specified	The value of a *purchase order* must be specified
The value of event data type *a1* is equal to the value of event data type *a2* and *a3*	The values of *Purchase order*, *goods receipt*, and *invoice receipt* must match before the corresponding invoice can be paid
The value of data type *a1* may not change before/at/ after a completion of activity *a2*	The value of a *purchase order* may not change after a *sign* activity has been performed

Note that the rule patterns presented in Table 19 are based on best practices for the P2P process and attempt to be comprehensive in mitigating risks associated with management assertions and fraudulent activities.

The Continuous Monitoring Layer

The continuous monitoring layer is based on the continuous assurance architecture put forward by Vasarhelyi et al. (2004). As previously mentioned, the continuous monitoring methodology relies on the business process to be formal or formalizable so that the information system can capture the data and feed automatically generated event log to the continuous monitoring layer (Alles et al. 2006). The design of the system architecture to continuously monitor the business process flow will be based on implementing a semantic abstracted layer placed on top of the business process, which stores the "ideal"

model for that business process (P2P in the case of this study) and the rules defined in the previous phase. The CM layer is based on an independent system called the monitoring and control layer (MCL) (Vasarhelyi et al. 2004). In a 3-tier architecture ERP system consisting of the presentation, application, and database layers, the MCL would interact with the application tier due to the complexity and enormity of the database tier (Alles et al. 2006). The MCL will rely on a read-only access of the event log data at the application layer.

As the steps of the transaction are being executed by the employees, the MCL will be comparing each step taken to the prescribed model and the defined rules associated with it to determine whether the transaction is violating any rules. For example, if an employee that created a purchase is attempting to release the same purchase order then the MCL would block the completion of the transaction since it is a violation of segregation of duties. Or, if a goods receipt activity gets inputted into the system before an authorization of the purchase order, then the MCL would either block it from being competed or allow it but flag it for further investigation.

This component of continuous assurance allows not only the detection of anomalies or errors, but also the prevention of completing an erroneous transaction. It ensures that routine transactions comply with prescribed process flows and controls. So, the ultimate objective is that no transaction would go through the system if it violates any of the rules set in place by management, such as limits, user access, or sequence of activity. This component requires imposing constraints on transaction activities at time of execution. Therefore, if an activity is in violation of a specific control, the system would intercept and block the execution of further activities in the transaction.

However, if there's an override for a certain violating activity, which could be allowable by management, then the system would allow the transaction to proceed while issuing an alert, and record the violation in an activity log maintained for audit purposes. This is essential because a sustainable approach for achieving compliance should fundamentally have a preventative focus, which is the objective of this component (Sadiq et al. 2007).

Additionally, information systems have the capabilities of configuring application controls in a way that would prevent many of the violation currently detected by process mining techniques or other audit analytics. However, this stringent configuration of controls might hinder the business operation of the firm since it would slow down or restrict many transactions from being executed. Therefore, continuous process mining assurance is a solution for this dilemma that would allow the business to be flexible and at the same time continuous monitor transactions for any errors or violations.

4.4. DEMONSTRATION OF METHODOLOGY

4.4.1. DATA

To demonstrate the viability of the framework presented in this study, an event log dataset that relate to the P2P process of a national non-profit professional organization is used. The reason for choosing the organization's P2P process is because it is a standardized typical business process similar to most businesses worldwide, which makes the study more generalizable. In addition, the procurement process represents a significant account that totals 73 million for 2016.

The event logs were constructed from the organization's ERP system that is used for

its procurement cycle. All the transactions used in this study consists of invoices billed during 2016, which are traced back to their accompanying purchase orders (PO)s that were created between June 2014 and December of 2016.

The entire population of data was used in the analysis because process mining techniques facilitates the option to do so. This demonstrates the ability of process mining to be used as an auditing tool that allows auditors to examine the population of data instead of resorting to sampling. In addition, since the entire population is examined, it becomes unnecessary to distinguish between analytical procedures and tests of detail (Jans et al. 2014).

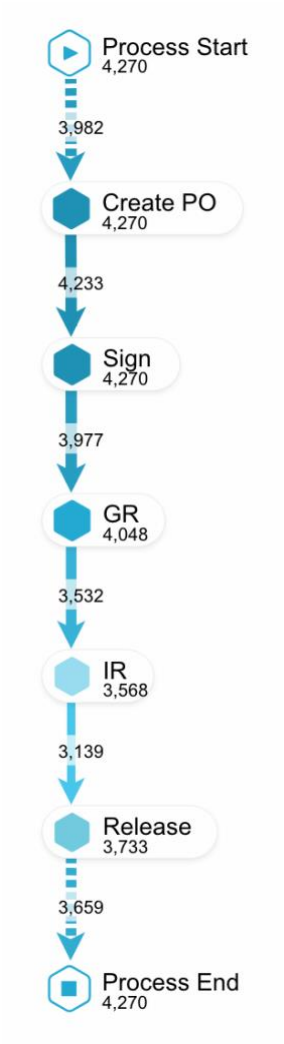
Furthermore, given the limitation of the data used to demonstrate the applicability of developing and implanting a continuous monitoring process mining layer for a business process (P2P), parts of the data was simulated based on the process model to cover the full business process, such as the roles of employees. Also, some deviations were simulated that represent close to real-life process deviations.

Next, analysis of the event log dataset and some examples of violated rules for a P2P process will be shown and how would the MCL prevent such violations:

4.4.2. PROCURE-TO-PAY BUSINESS PROCESS

Since the first step to implementing continuous monitoring using rule-based process mining techniques is to understand the underlying business process, which in this case is the P2P process, a process discovery analysis has been conducted to create a baseline model for routine transactions. The baseline model for the application is shown in Figure 17.

Figure 17. Standard Process Model for P2P Process



Process discovery allows auditors to reconstruct actual transaction workflows to verify any compliance issues and works as the starting step for defining and refining the rules that would continuously monitor transactions as they are being executed in the information system.

The effectiveness of controls in any business process depends on whether each of several multiple assertions are valid or not. Typical assertions to be considered are

“Existence”, “Completeness”, “Valuation” and “Presentation” (AU Sec. 326). Broadly speaking, for each management assertion, there are several potential risks, and for each risk there may be more than one internal control to mitigate the risk (Mock et al. 2009).

The P2P process tested in this study is designed based on business rules provided by the company and industry standards. This process begins with the creation of a Purchase Order (PO) from a valid vendor. The PO must be Signed. Once the PO is signed, the employee can order the goods or services from an authorized supplier. Then, the supplier will dispatch both the goods and the invoice related to it. Once that happens, both the Goods Receipt (GR) and Invoice Receipt (IR) are entered into the information system, and consequently, the accounts payable employee records the invoice in the general ledger. After that, an authorized personnel would release the PO which will initiate the Pay activity. The PO must be Signed and Released by two different and authorized employees, as a segregation of duties measure. Note that some deviations might occur to this design. Some of these deviations might be acceptable to the modeled design, such as goods being received over a period of time. Other deviations might be breakdowns in the internal control process, such as authorizing a payment without any GR or IR.

Just by conducting process discovery analysis, a large number of variants can be found. The P2P process had 1,061 unique sequence of activities in the event log and 4,270 cases. Hence, a rule-based process mining technique is essential and proves to be valuable in uncovering and preventing complex rule violations and deviations as transactions are being executed.

4.4.3. APPLICATION SCENARIOS

The following scenarios demonstrate the application of a continuous monitoring layer using rule-based process mining techniques to various financial transactions.

Acceptable Transaction Cycle

Routine and non-routine transactions should follow the business rules set forth by management. However, deviations from the standard model might still occur. These deviations need to be monitored to assure that they are acceptable. For example, case #87137 from the event log dataset had the following sequence:

Create PO —> Sign —> Sign —> GR —> IR —> GR —> IR —> Release

Even though this case had some repetitions in the GR and IR activities, it was acceptable and did not violate any business rules. Therefore, the continuous monitoring system would not flag or block this transaction and would allow it to complete its cycle. The same result would be to all transactions that follow the standard process model.

Suspicious Transaction Cycle

The set of rules have already defined that if a certain activity is performed then a different activity must be performed. In the case of this study, if a GR activity is performed then an IR activity must be performed. However, this is not always the case. For example, case # 89064 has a sequence of:

Create PO —> Sign —> Sign —> GR —> GR —> Release

The created PO value for this case was $8,000 (requires two signatures since it is above $5,000) and the goods were received in two installments over a span of four months ($5,500

in February 2016 and $2,500 in June 2016). However, this PO was released without an invoice being entered into the system, which violates one of the rule patterns. Not only that, but the accounts payable department was billed for that PO. Case # 89064 is not the only case that violates the specified rule pattern. There were 416 other cases that have the same control violation and account for 9% of the total cases in the dataset. This is a key concern that requires auditors' attention. The ERP system should have an indicator flagging IR for no system entry, but possibly this action had been overlooked by originators. An auditor would likely want to investigate these variants to ensure that the goods were actually received in total and the reason for releasing these POs without an IR. Therefore, had a continuous monitoring layer been implemented, it would prevent the release of the PO without the appropriate IR entered into the information system.

Another scenario that illustrates the benefits of process mining is being able to capture violations or anomalies that wouldn't have been caught when using traditional auditing methods. Case # 87130 has a sequence of:

Create PO —> Sign —> Sign —> GR —> IR —> GR —> IR —> GR —> IR —> GR —> IR —> Release —> Create PO —> GR —> IR —> Release

Even though this case followed normal business procedures for acquiring two signatures for the PO (it has a value of $48,750), it was released prior to all goods being received, which required it to create an additional PO activity to record the last GR. This violates the rule pattern that a certain activity (GR) must be completed before t time units of the final activity (Release). The continuous monitoring layer can prevent the first release of this transaction prior to the completion of GR. However, what is interesting about this case is that the following case (# 87131) is a duplication of the previous case with a slight change

in timestamp. Table 20 illustrates the duplication in the two cases 87130 & 87131.

Table 20. Duplicate POs with Slightly Different Timestamps

Activity	Originator	Timestamp	Value PO	Value GR	Value Pay
Create PO	P1	08/03/2015 13:18:48	48750.00		
Sign	P1	08/03/2015 13:18:49	48750.00		
Sign	P2	08/07/2015 09:57:05	48750.00		
GR	P3	11/04/2015 10:47:37	48750.00	11,250	
IR	P3	11/09/2015 14:12:12	48750.00	11,250	11250.00
GR	P3	02/04/2016 12:00:59	48750.00	12,500	
IR	P3	02/05/2016 14:58:30	48750.00	12,500	12500.00
GR	P3	05/09/2016 11:23:19	48750.00	12,500	
IR	P3	05/10/2016 10:33:15	48750.00	12,500	12500.00
Release	P4	07/01/2016 12:02:11	48750.00		
Create PO	P1	07/21/2016 09:25:06	48750.00		
GR	P3	07/21/2016 09:34:55	48750.00	12,500	
IR	P3	07/25/2016 16:29:14	48750.00	12,500	12500.00
Release	P5	07/26/2016 07:01:31	48750.00		
Activity	**Originator**	**Timestamp**	**Value PO**	**Value GR**	**Value Pay**
Create PO	P1	08/03/2015 13:22:37	48750.00		
Sign	P1	08/03/2015 13:22:38	48750.00		
Sign	P2	08/07/2015 09:57:18	48750.00		
GR	P3	11/04/2015 10:44:13	48750.00	11,250	
IR	P3	11/09/2015 14:14:02	48750.00	11,250	11250.00
GR	P3	02/04/2016 11:26:19	48750.00	12,500	
IR	P3	02/05/2016 14:55:38	48750.00	12,500	12500.00
GR	P3	05/09/2016 11:22:05	48750.00	12,500	
IR	P3	05/10/2016 10:39:15	48750.00	12,500	12500.00
Release	P4	07/01/2016 12:02:47	48750.00		
Create PO	P1	07/21/2016 09:24:43	48750.00		
GR	P3	07/21/2016 09:31:15	48750.00	12,500	
IR	P3	07/25/2016 16:30:39	48750.00	12,500	12500.00
Release	P5	07/26/2016 07:01:31	48750.00		

Even if these are sound transactions or that the system made an error in duplicating POs, the ERP system settings might allow them without flagging them for further investigation if they don't violate any of the settings. However, with a continuous monitoring rule-based process mining layer, these transactions would be flagged for suspicious purposes based on the fact that they are identical with a slight change in the timing of each activity.

A duplicate payment suspicion can also be found with other cases, such as case # 89501. In this case, the employee that entered the IR into the information system did it twice with two different timestamps, which lead the organization to be billed twice. Table 21 shows the sequence of the suspicious case.

Table 21. Duplicate Payment Due to Redundant Activity

Activity	Originator	Timestamp	Value PO	Value GR	Value Pay
Create PO	P1	02/12/2016 14:17:04	600.00		
Sign	P1	02/12/2016 14:17:05	600.00		
Sign	P2	02/16/2016 07:42:31	600.00		
GR	P3	02/16/2016 09:44:20	600.00	600	
IR	P3	02/17/2016 15:16:37	600.00	600	600.00
IR	P3	02/17/2016 15:17:49	600.00	600	600.00
Release	P4	02/18/2016 07:01:17	600.00		

The same employee was found to have performed the same duplicate activity in other cases. The continuous monitoring layer should intercept the second IR and payment before it is processed to prevent any intentional or unintentional errors.

Another rule pattern that is tested against the dataset is that a certain activity must be performed at least once. Case # 90979 violates this rule by not having a Release activity performed after GR and IR. This case has a sequence of:

Create PO —> Sign —> Sign —> GR —> IR

This PO was billed without having a release activity that sends it to the accounts payable department. This would be an opportunity for the continuous monitoring layer to block payment to the PO without proper Release activity associated with it.

A similar case that also is missing key activities (IR and Release) is case # 89092, which has the following sequence:

Create PO —> Sign —> Sign —> GR —> GR —> Create PO

This case violates many rules including the duplication of the GR activity by two different employees, and ending the transaction with a Create PO instead of IR and Releasing the PO. So, the continuous monitoring layer would flag the second GR activity and would not allow the transaction to end on a PO creation. It would require the auditor to further investigate this transaction.

Another control that can be found in ERP systems is 3-way match. This control is configured so that when a good is received, a note is entered into the system. Then the system checks it against the PO. When an invoice is entered into the system, the system cross checks the details to the PO and, where applicable, to the GR note. So, if the details agree, or within tolerance, the invoice is approved for payment. However, if the details do

not agree, the invoice is listed on an exception report, for manual follow-up. Invoices should not be paid until they have been cleared from the exception report.

However, there are cases in which the value of the PO does not match the value of GR (Case # 89092, for example). In this case, if the PO was not entered correctly or there was an error in entering the GR, and the 3-way match embedded in the ERP system did not flag it or list the PO in the exceptional report, then an implemented continuous monitoring layer would flag it and prevent payment.

Additionally, a PO with a value of $64.71 in case # 89554 had an extremely high payment made to it with an amount of $21,783.05. This payment does not seem to be from a routine transaction. However, these anomalistic transactions should be further investigated to assure that they are approved transactions or that the system is not making errors in payments or recording them.

What is interesting also is that it was found that the same employee (resource) that created the PO, GR, and IR of the last case also had many other suspicious transactions. For example, cases # 87134 and 87135 had the same employee create the PO, Sign, GR, and IR. And the payment made against the PO for case # 87134, which had a value of $225.55, was $50,000.53. Also, the payment made against the PO for case # 87135, which had a value of $1,710.35, was $50,000.53. These transactions raise many flags for segregation of duties controls and suspicious payments. Process mining allows for uncovering these suspicious activities and the information gained from it would be very valuable if it was timely.

Continuous monitoring using process mining techniques would uncover a string of suspicious activities. For example, case # 88893 had also a large payment ($56,099.08)

made against a relatively low value PO ($750). This case had a normal sequence:

Create PO —> Sign —> GR —> IR —> Release

The large payment should flag this transaction for further investigation. However, there was a trend of 17 cases that had the exact same amount of payment made against. Hence, the continuous monitoring layer would implement a rule that would monitor payment if there is a certain repetition in terms of the amount for more than 5 cases for example, any subsequent transaction would be flagged for further investigation.

There is a business rule specified that a Create PO activity must be started before date of IR. But, analysis of the event log shows that there are 6 cases where the first activity is an IR not a PO creation. And two of the cases also had a suspicious large payment. Table 22 shows the sequence of the two cases and value (87756 & 87823)

Table 22. Suspicious Activities in Incorrect Order

Activity	Originator	Timestamp	Value PO	Value GR	Value Pay
IR	P1	09/18/2015 14:50:38	46,028.35	46,028.35	113,528.35
Create PO	P2	09/21/2015 11:27:29	46,028.35		
Sign	P2	09/21/2015 11:27:30	46,028.35		
Sign	P3	09/21/2015 12:32:19	46,028.35		
GR	P1	09/21/2015 12:54:43	46,028.35	46,028.35	
Release	P4	09/22/2015 13:43:23	46,028.35		

Activity	Originator	Timestamp	Value PO	Value GR	Value Pay
IR	P1	09/25/2015 08:47:54	145.00	145.00	25,597.02
Create PO	P1	09/25/2015 09:33:13	145.00		
Sign	P1	09/25/2015 09:33:14	145.00		
GR	P1	09/25/2015 09:33:31	145.00	145.00	
Release	P2	09/25/2015 09:56:13	145.00		

This sequence wouldn't be allowed had a continuous monitoring layer been implemented to intercept these violations of business rules.

There is also segregation of duties rules that can be configured to be detected by the continuous monitoring layer. The rule pattern that a person must not perform both activities for two separate roles needs to be implemented. However, this was not the case in the event log. For example, case # 87712 has the same user perform the activities of Create PO, Sign, and Release. This is a violation of segregation of duties. Additionally, there are other cases that have the same employee Create PO, Sign, GR, IR (Case # 87216 for example).

The continuous monitoring layer can not only be for detecting and preventing erroneous transactions but also for other issues such as improving efficiency. For example, if a PO has a value less then $5,000, then it does not require two signatures to be released. But, there are cases where there are multiple signatures for a very low PO amount. Case # 87696 is a good example since it has a PO value of $14 and yet had two signatures. This inefficiency needs to be improved and the continuous monitoring layer can help achieve that by notifying the employees about the business rules.

Another issue is providing timely information from process mining, which can benefit a company by reducing the time it takes to release a PO once all the required previous steps had been completed. For example, case # 87287 was released more than a month after the last IR was entered, which could cost the company for late payments. This issue needs to be addressed to improve efficiency and reduce cost. Table 23 shows the details of the case.

Table 23. Unacceptable Amount of Time for Releasing a PO

Activity	Originator	Timestamp	Value PO	Value GR	Value Pay
Create PO	P1	06/10/2014 14:10:15	2375.00		
Sign	P1	06/10/2014 14:10:15	2375.00		
Sign	P2	06/10/2014 14:57:58	2375.00		
Create PO	P3	08/12/2015 14:06:43	2375.00		
Sign	P3	08/12/2015 14:06:43	2375.00		
Sign	P4	08/17/2015 17:20:27	2375.00		
GR	P5	09/02/2015 11:11:15	2375.00	475.00	
IR	P5	09/04/2015 09:16:29	2375.00	475.00	475.00
GR	P5	09/24/2015 08:08:04	2375.00	475.00	
IR	P5	09/28/2015 08:15:17	2375.00	475.00	475.00
GR	P5	05/04/2016 09:28:49	2375.00	475.00	
IR	P5	05/06/2016 08:46:37	2375.00	475.00	475.00
GR	P5	05/26/2016 10:03:27	2375.00	475.00	
GR	P5	06/22/2016 10:10:20	2375.00	475.00	
IR	P5	06/26/2016 14:16:59	2375.00	475.00	475.00
IR	P5	08/10/2016 14:50:38	2375.00	475.00	475.00
Release	P6	09/30/2016 15:13:08	2375.00		

4.5. CONCLUSION

This paper proposes a novel approach for assurance that utilizes the advantages of a continuous monitoring layer using rule-based process mining techniques. Rather than reacting after the violations have long occurred, this solution allows auditors to actively detect and investigate deviations and exceptions as they occur along the transaction process by continuously monitoring business process controls and testing transactions. Any transaction that violates a set of business rules would be intercepted or flagged by the system until investigated by an auditor. This approach provides a high level of assurance on the operating effectiveness of controls throughout a business process. The framework was applied to a specific area in the procurement business process to showcase how it

would function. However, the data had its limitations since some controls that would be required for that business process had to be simulated, such as roles of employees. It would be interesting to apply this model to other business cycles or a richer dataset.

Besides the limitation of the data, implementing a continuous monitoring layer using a rule-based process technique has its challenges. Firstly, the successful implementation depends on the underlying technology of the information system. The information system has to be "process aware" and the logging capabilities turned on to allow for the automatic creation of the event log. Not only that, but the attributes included in the event log can affect the effectiveness of the continuous monitoring layer and its ability to identify and capture violations. Additionally, the extensiveness of this continuous assurance solution depends on the comprehensiveness of the rules imbedded in the continuous monitoring layer. Hence, the exhaustiveness of the rules will determine how effective this continuous assurance is.

The use of process mining as a preventative approach rather than detective is rarely found in the auditing literature. As such, it contributes to the sparse literature on internal controls effectiveness assessment and compliance. In future work, one can utilize different weighting schemas and various variables, such as the number of controls violated or the cost of violated transaction, to provide a hierarchal listing of flagged transactions to be furtherly investigated by auditors. Also, a real-life pilot implementation of this framework would be the focus of future work to assess the feasibility and accuracy of such methodology.

CHAPTER 5: CONCLUSION

This dissertation contributes to the auditing literature by exploring the application of process mining and its evolution in auditing, specifically in internal controls. Process mining is a tool that can be used by auditors to help with discovering, monitoring, and improving actual processes by extracting knowledge from unstructured data sources. Implementing process mining in the audit process allows auditors to test the entire population, rather than a sample in the traditional setting, and base their opinions on an objective data source in the form of meta-data from the company's ERP system.

Process mining has the ability to substantially improve the audit process in terms of its effectiveness and efficiency. By examining the entire population of data that cannot be altered or manipulated with, auditors can gather significantly strong evidence to support their findings and professional judgment. Even though process mining is not without its shortcomings (mostly its dependence on event log data that is usually not easy to collect), its benefits drastically outweigh its limitations, and hence, auditors and standard setters must make it part of the audit process. This dissertation demonstrates in three essays how process mining can improve the audit process in different aspects.

The first essay develops a methodology for objectively measuring the effectiveness of internal controls and risk assessment. So, instead of relying on traditional and qualitative methods, the general framework would provide auditors with a more objective and efficient way of assessing if controls are implemented, and to what degree. The methodology developed in the first essay illustrates how process mining can be used to test internal controls to provide an overall risk assessment of the internal control system for a business

process. This essay answers the call in the auditing literature for developing a baseline of control effectiveness measurement. Regulations require auditors to assess internal control risk both in terms of implementation and operation. In a traditional audit, auditors rely on the use of sampling due to the labor and time intensiveness of manual testing. In contrast, advanced audit tools, such as process mining, would consider the whole population of transactions in testing. The consideration of the whole population of transactions in testing can enhance the effectiveness of an audit and increases the probability that material errors, omissions, fraud, and internal control violations may be detected.

In the first essay, the conceptual model was tested on a set of data that relates to the procurement process obtained from a national not-for-profit organization. The results have found several internal controls to be lacking in different areas of the procurement process. The results of the analysis were visualized on a risk map where the violations and their impact provide auditors with an instant way of highlighting areas of increased risk.

This essay had its limitations. First, some controls needed to be simulated since the firm that provided the data did not provide feedback on what control were implanted. So, the study had to rely on industry standard and the literature to come up with the controls that were tested. Second, some of the assertions tested in this study, such as completeness, cannot by fully examined using process mining, since it relies on the logs that were extracted from the information system of the firm. It is strongly assumed that the ability to manipulate the event logs is not present. Finally, there are concerns with calculating the effectiveness of the internal control system for a business process with controls that are not included in the testing or the scores calculated. These controls would include manual, but essential, controls but are not captured in the event log found in the information system.

In future research, one can utilize different weighting schemas and various variables to enhance the calculation method and provide a more accurate measure for risk assessment. Additionally, research should be done on how the materiality principle changes with process mining analysis and full population testing. Lastly, applying this methodology on other business process and organization would help in its generalizability.

The second essay is concerned with a challenge that auditors face in finding ways to detect and investigate anomalies and exceptions with the large number of transactions being executed on a daily basis. Therefore, the aim and contribution of the second essay is a methodology that provides auditors with guidance on the use of process mining in conjunction with existing analytical procedures to identify exceptional transactions that would require further investigation. This solution allows auditors to focus on process instances that are likely to be considered high-risk, reduce the risk of failing to detect material misstatement, and enhance audit effectiveness. Furthermore, the identification and prioritization of such risky process instances help with the information overload problem that entails population testing.

The process instances risk prioritization framework is based on applying process mining techniques on an event log extracted from the organization's information system to detect anomalies. These anomalies are then filtered using other analytical procedures to identify high-risk exceptional process instances. The exceptional process instances are then prioritized based on a calculated risk score. This combination of process mining with other analytical procedures is unique to this study.

The framework proposed in this study was demonstrated on a real-life event log

dataset that was obtained from the procure-to-pay process of a not-for-profit national organization. The event log contained a total of 4,142 process instances. After applying the first part of the framework, which is the process mining part, it highlighted 1,346 notable process instances. Existing analytical procedures were then applied, which is the second part of the framework, this resulted in narrowing down the results of problematic process instances to 814 exceptional process instances. A threshold was then applied to focus on process instances with a monetary value above a certain amount, which resulted in highlighting only 457 highly problematic process instances that have a material amount.

The limitations associated with this essay is that the results of exceptional process instances could differ depending on the filters developed and the weights given to each filter. Also, the framework was illustrated using an event log from a business process of one organization, and therefore, the conclusions drawn could be somewhat limited. It would be optimal if methodology could be implemented in a real-world audit sampling environment to test its effectiveness and efficiency. Future research could accomplish that.

Even though this study applies process mining prior to other existing analytical procedures, future studies could examine whether the prioritized exceptional process instances would differ if process mining is applied in the final stages of the framework as opposed to the early stages and compare the results. Furthermore, future work can include the implementation of the risk score methodology found in this study to a continuous process mining solution as a way to allow or block transactions based on their risk score.

The third essay conceptualizes the evolution of process mining by proposing a novel approach for assurance that utilizes the advantages of a continuous monitoring layer

using rule-based process mining techniques. By significantly reducing the time-delay, the created information becomes more valuable since it allows for additional management control and assurance activities.

This essay contributes to the auditing literature by developing a novel approach for monitoring assurance that combines the advantages of continuous monitoring with those of process mining. Auditors can actively detect and investigate deviations and exceptions as they occur along the transaction process by continuously monitoring business process controls and testing transactions, rather than react after the exceptions have long occurred. Any transaction that violates a set of business rules would be intercepted or flagged by the system until investigated by an auditor. This continuous monitoring using rule-based process mining approach provides a high level of assurance about the operating effectiveness of controls throughout a business process.

The conceptual framework was demonstrated on a procurement process using event logs obtained from a national non-profit organization to showcase how it would function. However, the data had its limitations since some rules and controls that would be required for that business process had to be simulated, such as roles of employees. It would be interesting to apply this model to other business cycles or a richer dataset.

Besides the limitation of the data, implementing a continuous monitoring layer using a rule-based process technique has its challenges. Firstly, the successful implementation depends on the underlying technology of the information system. The information system has to be “process aware” and the logging capabilities turned on to allow for the automatic creation of the event log. Not only that, but the attributes included

in the event log can affect the effectiveness of the continuous monitoring layer and its ability to identify and capture violations. Additionally, the extensiveness of this continuous assurance solution depends on the comprehensiveness of the rules imbedded in the continuous monitoring layer. Hence, the exhaustiveness of the rules will determine how effective this continuous assurance is.

The use of process mining as a preventative approach rather than detective is rarely found in the auditing literature. As such, it contributes to the sparse literature on internal controls effectiveness assessment and compliance. In future work, one can utilize different weighting schemas and various variables, such as the number of controls violated or the cost of violated transaction, to provide a hierarchal listing of flagged transactions to be furtherly investigated by auditors. Also, a real-life pilot implementation of this framework would be the focus of future work to assess the feasibility and accuracy of such methodology.

REFERENCES

Aalst, W. van der (2011). Process mining: Discovery, conformance and enhancement of business processes.

Abraham, A., (2005). Rule-based Expert Systems. Handbook of Measuring System Design, John Wiley & Sons, Ltd, 909-919(8)

Act, Sarbanes-Oxley. "Public Law No. 107-204." *Washington, DC: Government Printing Office* 107 (2002)

Adriansyah, A., van Dongen, B. F., & van der Aalst, W. M. (2011, August). Conformance checking using cost-based fitness analysis. In *Enterprise Distributed Object Computing Conference (EDOC), 2011 15th IEEE International* (pp. 55-64). IEEE.Agrawal, R., Gunopulos, D., & Leymann, F. (1998, March). Mining process models from workflow logs. In *International Conference on Extending Database Technology* (pp. 467-483). Springer, Berlin, Heidelberg.

Agrawal, R., Johnson, C., Kiernan, J., & Leymann, F. (2006). Taming Compliance with Sarbanes-Oxley Internal Controls Using Database Technology. In (Liu, L.; Reuter, A.; Whang, K.-Y.; Zhang, J. Hrsg.): Proc. In *22nd International Conference on Data Engineering (ICDE 2006), Atlanta* (p. 92).

Alles, M. G., Kogan, A., & Vasarhelyi, M. A. (2002). Feasibility and economics of continuous assurance. *Auditing: A Journal of Practice & Theory*, *21*(1), 125-138.

Alles, M., Kogan, A., & Vasarhelyi, M. (2003). Black box logging and tertiary monitoring of continuous assurance systems. *Information Systems Control Journal*, *1*, 37-41.

Alles, M. G., Kogan, A., & Vasarhelyi, M. A. (2004). Restoring auditor credibility: tertiary monitoring and logging of continuous assurance systems. *International Journal of Accounting Information Systems*, *5*(2), 183-202.

Alles, M., Brennan, G., Kogan, A., & Vasarhelyi, M. A. (2006). Continuous monitoring of business process controls: A pilot implementation of a continuous auditing

system at Siemens. *International Journal of Accounting Information Systems*, *7*(2), 137-161.

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

Amat, J. L. (2002). Using reporting and data mining techniques to improve knowledge of subscribers; applications to customer profiling and fraud management. *Journal of Telecommunications and Information Technology*, 11-16.

American Institute of Certified Public Accountants. Auditing Standards Board. (1983). *Audit risk and materiality in conducting an audit*. The Institute.

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

American Institute of Certified Public Accountants (AICPA). 1997. *The Information Technology Age: Evidential Matter in the Electronic Environment.* Jersey City, NJ: AICPA.

Aobdia, D., Siddiqui, S., & Vinelli, A. G. (2016). Does engagement partner perceived expertise matter? Evidence from the US operations of the Big 4 audit firms.

Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017). Big Data and analytics in the modern audit engagement: Research needs. *Auditing: A Journal of Practice & Theory*, *36*(4), 1-27.

Arens, A. A., Elder, R. J., & Mark, B. (2012). *Auditing and assurance services: an integrated approach*. Boston: Prentice Hall.

Arens, A. A., Loebbecke, J. K., Elder, R. J., Beasley, M. S., & American Institute of Certified Public Accountants. (2000). *Auditing: An integrated approach* (Vol. 8). Upper Saddle River, NJ: Prentice Hall.

Ashbaugh-Skaife, H., Collins, D. W., & Kinney, W. R. (2007). The discovery and reporting

of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics*, *44*(1), 166-192.

Azzini, A. & Damiani, E. (2015). Process mining in big data scenario. In Proceedings of SIMPDA 2015, volume 1527, pages 149–153.

Bailey Jr, A. D., Duke, G. L., Gerlach, J., Ko, C. E., Meservy, R. D., & Whinston, A. B. (1985). TICOM and the analysis of internal controls. *Accounting Review*, 186-201.

Bierstaker, J. L., & Wright, A. (2004). Does the adoption of a business risk audit approach change internal control documentation and testing practices?.*International Journal of Auditing*, *8*(1), 67-78.

Board, I. S. (2002). Continuous auditing: Is it fantasy or reality?. *Information Systems Control Journal*, *5*.

Borthick, A. F. (2012). Designing continuous auditing for a highly automated procure-to-pay process. *Journal of Information Systems*, *26*(2), 153-166.

Brewster, B. (2008). *Enhancing the auditor expertise model: How systems thinking fosters a reinforcing feedback loop between knowledge and ability*. Working paper, University of Illinois.

Carnaghan, C. (2006). Business process modeling approaches in the context of process level audit risk assessment: An analysis and comparison. *International Journal of Accounting Information Systems*, *7*(2), 170-204.

Caron, F., & Vanthienen, J. (2012). Applications of business process analytics and mining for internal control. *ISACA Journal: the source of IT governance professionals*, *4*, 44-49.

Caron, F., Vanthienen, J., & Baesens, B. (2013). A comprehensive investigation of the applicability of process mining techniques for enterprise risk management. *Computers in Industry*, *64*(4), 464-475.

Caron, F., Vanthienen, J., & Baesens, B. (2013). Comprehensive rule-based compliance checking and risk management with process mining. *Decision Support Systems*, *54*(3), 1357-1369.

Chan, K. C., Farrell, B., & Lee, P. (2008). Earnings management of firms reporting material internal control weaknesses under Section 404 of the Sarbanes-Oxley Act. *Auditing: A Journal of Practice & Theory*, *27*(2), 161-179.

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

Chang, I. C., & Liu, C. C. (2013). Assessment Mechanism of Internal Control for Information Technology Governance. *Assessment*, *6*, 18-2013.

Chang, S. I., Yen, D. C., Chang, I. C., & Jan, D. (2014). Internal control framework for a compliant ERP system. *Information & Management*, *51*(2), 187-205.

Changchit, C., Holsapple, C. W., & Madden, D. L. (2001). Supporting managers' internal control evaluations: an expert system and experimental results.*Decision Support Systems*, *30*(4), 437-449.

Chiu, T., Vasarhelyi, M., Alrefai, A., & Yan, Z. (2018). Validating Process Mining: A Framework Integrating Auditor's Risk Assessment.

Cooley, J. W., & Cooley, B. J. (1982). Internal accounting control systems: A simulation program for assessing their reliabilities. *Simulation & Games*, *13*(2), 211-231.

Cushing B. E. (1974), "A Mathematical Approach to the Analysis and Design of Internal Control Systems", *The Accounting Review,* Vol. 49, No. 1, pp. 24-41.

Cushing B. E. (1975), "A Further Note on the Mathematical Approach to Internal Control", *The Accounting Review*, Vol. 50, No. 1, pp. 151-154.

Datta, A. (1998). Automating the discovery of as-is business process models: Probabilistic and algorithmic approaches. *Information Systems Research*, *9*(3), 275-301.

Davenport, T. H. (1993). *Process innovation: reengineering work through information*

technology. Harvard Business Press.

Debreceny, R., Gray, G. L., Tham, W. L., Goh, K. Y., & Tang, P. L. (2003). The development of embedded audit modules to support continuous monitoring in the electronic commerce environment. *International Journal of Auditing*, *7*(2), 169-185.

Elliott, R. K. (2002). Twenty-first century assurance. *Auditing: A Journal of Practice & Theory*, *21*(1), 139-146.

Ernst & Young LLP. (2002), "Preparing for Internal Control Reporting: A Guide for Management's Assessment Under Section 404 of the Sarbanes-Oxley Act".

Graham, L. E. (1993). Discussion of expertise in auditing. *Auditing: A Journal of Practice & Theory*, *12*, 46-50.

Grimlund, R. A. (1982). An integration of internal control system and account balance evidence. *Journal of Accounting Research*, 316-342.

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

Günther, C. W., & Van Der Aalst, W. M. (2007, September). Fuzzy mining–adaptive process simplification based on multi-perspective metrics. In *International Conference on Business Process Management* (pp. 328-343). Springer Berlin Heidelberg.

Hakvoort, R., & Sluiter, A. (2008). Process Mining: Conformance analysis from a financial audit perspective. *Int. J. Business Process Integration and Management*.

Hammer, M., & Champy, J. (1993). Reengineering the corporations.

Hosseinpour, M., & Jans, M. (2016). Categorizing Identified Deviations for Auditing. In *SIMPDA* (pp. 125-129).

Hwang, S. S., Shin, T., & Han, I. (2004). CRAS-CBR: Internal control risk assessment

system using case-based reasoning. *Expert Systems*, *21*(1), 22-33.

Islam, A., Corney, M., Mohay, G., Clark, A., Bracher, S., Raub, T., and Flegel, U. (2010). Fraud detection in ERP systems using scenario matching. *Security and Privacy–Silver Linings in the Cloud*, pages 112–123.

Issa, H. (2013). *Exceptional exceptions* (Doctoral dissertation, Rutgers University-Graduate School-Newark).

Issa, H., & Kogan, A. (2014). A predictive ordered logistic regression model as a tool for quality review of control risk assessments. *Journal of Information Systems*, *28*(2), 209-229.

Jans, M. J. (2011, August). Process mining in auditing: From current limitations to future challenges. In *International Conference on Business Process Management* (pp. 394-397). Springer, Berlin, Heidelberg.

Jans, M., Alles, M., & Vasarhelyi, M. (2013). The case for process mining in auditing: Sources of value added and areas of application. *International Journal of Accounting Information Systems*, *14*(1), 1-20.

Jans, M., Alles, M. G., & Vasarhelyi, M. A. (2014). A field study on the use of process mining of event logs as an analytical procedure in auditing. *The Accounting Review*, *89*(5), 1751-1773.

Jans, M., Depaire, B., & Vanhoof, K. (2011). Does process mining add to internal auditing? an experience report. In *Enterprise, Business-Process and Information Systems Modeling* (pp. 31-45). Springer, Berlin, Heidelberg.

Jans, M., van der Werf, J. M., Lybaert, N., & Vanhoof, K. (2011). A business process mining application for internal transaction fraud mitigation. *Expert Systems with Applications*, *38*(10), 13351-13359.

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

of Emerging Technologies in Accounting, *9*(1), 95-110.

Kogan, A., Alles, M. G., Vasarhelyi, M. A., & Wu, J. (2010). Analytical Procedures for Continuous Data Level Auditing: Continuity Equations 1.

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

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

Krishnan, R., Peters, J., Padman, R., & Kaplan, D. (2005). On data reliability assessment in accounting information systems. *Information Systems Research*, *16*(3), 307-326.

Lavigne, A. (2003). *Electronic Audit Evidence.* Toronto, Canada: Canadian Institute of Chartered Accountants.

Leymann, F., & Altenhuber, W. (1994). Managing business processes as an information resource. *IBM systems journal*, *33*(2), 326-348.

Li, C., Reichert, M., & Wombacher, A. (2008, July). Mining process variants: Goals and issues. In *Services Computing, 2008. SCC'08. IEEE International Conference on* (Vol. 2, pp. 573-576). IEEE.

Li, P., D. Y. Chan, and A. Kogan. 2016. Exception prioritization in the continuous auditing environment: A framework and experimental evaluation. *Journal of Information Systems* 30 (2): 135-157.

Mautz, R. K., & Sharaf, H. A. (1961). *The philosophy of auditing* (No. 6). American Accounting Association.

Mock, T. J., Sun, L., Srivastava, R. P., & Vasarhelyi, M. (2009). An evidential reasoning approach to Sarbanes-Oxley mandated internal control risk assessment. *International Journal of Accounting Information Systems*, *10*(2), 65-

78.

Mock T., and Turner J. (1981), Internal accounting control evaluation and auditor judgment, American Institue of Certified Public Accountants, New York, NY, USA.

Namiri, K., & Stojanovic, N. (2007). A formal approach for internal controls compliance in business processes. In *8th Workshop on business process modeling, development, and support* (pp. 1-9).

Norman, C. S., Payne, M. D., & Vendrzyk, V. P. (2009). Assessing information technology general control risk: An instructional case. *Issues in Accounting Education Teaching Notes*, *24*(1), 25-43.

Public Company Accounting Oversight Board (PCAOB). (2007). An Audit of Internal Control over Financial Reporting That Is Integrated with an Audit of Financial Statements. Auditing Standard No. 5.

Public Company Accounting Oversight Board (PCAOB). 2010. Audit Evidence. Auditing Standard No. 15, PCAOB Release No. 2010-004. Washington, DC: PCAOB.

Rezaee, Z., Elam, R., & Sharbatoghlie, A. (2001). Continuous auditing: the audit of the future. *Managerial Auditing Journal*, *16*(3), 150-158.

Rezaee, Z., Sharbatoghlie, A., Elam, R., & McMickle, P. L. (2002). Continuous auditing: Building automated auditing capability. *Auditing: A Journal of Practice & Theory*, *21*(1), 147-163.

Rikhardsson, P., & Kræmmergaard, P. (2006). Identifying the impacts of enterprise system implementation and use: Examples from Denmark. *International Journal of Accounting Information Systems*, *7*(1), 36-49.

Rozinat, A., & Van der Aalst, W. M. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, *33*(1), 64-95.

Sadiq, S., Governatori, G., & Namiri, K. (2007, September). Modeling control objectives for business process compliance. In *International conference on business process management* (pp. 149-164). Springer Berlin Heidelberg.

Securities and Exchange Commission (SEC). 1941. *Amendment of Rules 2-02 and 3-07 of Regulation S-X*. Accounting Series Release No. 21, 11. Fed. Reg. 10921. February, 5.

Selig, H. (2017). Continuous Event Log Extraction for Process Mining.

Srivastava, R. P. (1986). Auditing functions for internal control systems with interdependent documents and channels. *Journal of Accounting Research*, 422-426.

Srivastava, R. P., & Ward, B. H. (1983, December). Reliability modeling of information systems with human elements: A new perspective. In *IEEE Transactions: Total Systems Reliability Symposium* (pp. 30-39).

The Committee of Sponsoring Organizations of the Treadway Commission (COSO). (1992), *Internal Control-Integrated Framework.*

The Committee of Sponsoring Organizations of the Treadway Commission (COSO). (2013), *Internal Control-Integrated Framework.*

Tuttle, B., & Vandervelde, S. D. (2007). An empirical examination of CobiT as an internal control framework for information technology. *International Journal of Accounting Information Systems*, *8*(4), 240-263.

Van der Aalst, W. M. (2009). Process-aware information systems: Lessons to be learned from process mining. In *Transactions on petri nets and other models of concurrency II*(pp. 1-26). Springer, Berlin, Heidelberg.

van de Aalst, W. (2010). Process discovery: capturing the invisible. *IEEE Computational Intelligence Magazine, 5*(1), 28-41.

Van der Aalst, W. M., & de Medeiros, A. K. A. (2005). Process mining and security:

Detecting anomalous process executions and checking process conformance. *Electronic Notes in Theoretical Computer Science*, *121*, 3-21.

van der Aalst, W. M., Reijers, H. A., Weijters, A. J., van Dongen, B. F., De Medeiros, A. A., Song, M., & Verbeek, H. M. W. (2007). Business process mining: An industrial application. *Information Systems*, *32*(5), 713-732.

Van der Aalst, W. M., van Dongen, B. F., Herbst, J., Maruster, L., Schimm, G., & Weijters, A. J. (2003). Workflow mining: A survey of issues and approaches. *Data & knowledge engineering*, *47*(2), 237-267.

van Aalst, W. M., van Hee, K. M., van Werf, J. M., & Verdonk, M. (2010). Auditing 2.0: Using process mining to support tomorrow's auditor. *Computer*, *43*(3).

Vasarhelyi, M. A., Alles, M. G., & Kogan, A. (2004). Principles of analytic monitoring for continuous assurance. *Journal of emerging technologies in accounting*, *1*(1), 1-21.

Vasarhelyi, M., & Greenstein, M. (2003). Underlying principles of the electronization of business: A research agenda. *International Journal of Accounting Information Systems*, *4*(1), 1-25.

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

Vasarhelyi, M. A., & Halper, F. B. (2002). Concepts in continuous assurance. *Researching accounting as an information systems discipline*, 257-271.

Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big Data in accounting: An overview. *Accounting Horizons*, *29*(2), 381-396.

Williamson, A. L. (1997). The implications of electronic evidence. *Journal of Accountancy*, *183*(2), 69.

Woodroof, J., & Searcy, D. (2001, January). Continuous audit implications of Internet technology: Triggering agents over the Web in the domain of debt covenant

compliance. In *System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on* (pp. 8-pp). IEEE.

Wu, J. H., Shin, S. S., & Heng, M. S. (2007). A methodology for ERP misfit analysis. *Information & Management*, *44*(8), 666-680.

Yu S. and Neter J. (1973), "A Stochastic Model of the Internal Control System", *Journal of Accounting Research,* Vol. 11, No. 2, pp. 273-295.