Public Auditing, Analytics, and Big Data in the Modern Economy

By

Deniz Appelbaum

A Dissertation submitted to the

Graduate School - Newark

Rutgers, the State University of New Jersey

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Professor Miklos A.Vasarhelyi

and approved by

____
Dr. Miklos A. Vasarhelyi

____
Dr. Alexander Kogan

____
Dr. Helen Brown-Liburd

____
Dr. Theodore J. Mock

Newark, New Jersey

May, 2017

Copyright page:

ABSTRACT OF THE DISSERTATION

Public Auditing, Analytics, and Big Data in the Modern Economy

by Deniz Appelbaum

Dissertation Director:

Professor Miklos A. Vasarhelyi

There is an increasing recognition in the public audit profession that the emergence of big data as well as the growing use of analytics by audit clients has brought new concerns and opportunities.

The first chapter introduces and identifies a number of these issues as that are facing the auditor in the modern economy.

The second chapter primarily addresses one of these concern: what is the extant research on analytical procedures in the audit engagement? This disertation proposes that the answers to these issues should start with an examination of the extant external audit research. However, an updated review of this research does not exist. Accordingly, 301 papers are identified regarding analytical procedures in the audit engagement. These papers are organized by technique, audit phase, and other attributes for understanding. This analysis is then presented as an External Audit Analytics (EAA) framework, which is subsequently expanded with the concepts of business analytics. Specifically, this synthesis organizes this literature, thereby

offering guidelines regarding possible approaches for more complex and data driven analytics in the engagement.

The third chapter elaborates and expands upon the next six issues and discusses additional aspects for contemplation by researchers and the profession.

The fourth chapter discusses the issues of Big Data when it is being considered as Audit Evidence, particularly in the context of external big data. In this age of big data many sources of evidence are untraceable and their provenance unverifiable. This chapter provides guidance regarding provenance of big data, allowing it to be regarded as reliable evidence for external auditors. Finally, the fifth chapter concludes.

These chapters discuss and illuminate broadly many issues facing the profession since clients are more automated and are capturing more data. These chapters also contribute to audit literature regarding external audit analytics and reliability of big data audit evidence.

Big data and analytics are dramatically changing the business environment and their processes. Business methods are changing, capabilities are being added, anachronistic functions are being eliminated, and processes are being substantially accelerated. The same paradigm change should occur with the audit profession, and this dissertation provides some of the needed ideas to motivate such a shift.

PREFACE: ACKNOWLEDGEMENT PAGE

This dissertation represents a huge accomplishment in my life, and for which I am indebted to many individuals for their support and encouragement. First, without the support from my husband Marc Appelbaum and my parents, this dissertation would not have happened. Thanks also to my dissertation committee members for their advice and constant and generous encouragement.

Likewise regarding the mentorship and guidance from the faculty here at Rutgers – Dr. Vasarhelyi, Dr. Kogan, Dr. Brown-Liburd, Dr. Palmon, Dr. Issa, Dr. Moffitt, Dr. Alles, Dr. Xiong, and Dr. Adams. I hope that our work together has just begun! I am also grateful to my colleagues Stephen Kozlowski, J.P. Krahel, Abdullah Alawhadi, Desi Arisandi, and Sophia Sung. I am also indebted to mentors and colleagues outside of the program – Ted Mock, Raj Srivastava, Rob Nehmer, Graham Gal, Brigitte Mueller, Glenn Grey, and Edson Riccio. Without our collaborative dicsussions, mutual papers, and research projects, this program would not have been such a rich and rewarding experience.

Finally, I feel compelled to extend a special acknowledgement to my advisor Dr. Vasarhelyi. Dr. Vasarehlyi inspired me to look at auditing and accounting in a totally different light and awakened my then-dormant passion for research. I have been blessed beyond measure to have had the opportunity to learn under his guidance and

tutelage. Very few individuals have the opportunity to study with such a great mind and inspirational thinker as Dr. Vasarhelyi. I will never forget the day when I first walked into his Information Technology course that first summer as a PAMBA student! Thank you Dr. Vasarhelyi for your belief in my ability to become a researcher of AIS.

TABLE OF CONTENTS

LIST OF TABLES:

LIST OF FIGURES:

CHAPTER ONE: INTRODUCTION

There is an increasing recognition in the external audit profession that the emergence of big data (Vasarhelyi, Kogan, and Tuttle 2015) as well as growing use of data analytics in business processes has brought a set of new concerns and opportunities to the audit community. It could be said that there are two forces that will exert huge impact on the profession: business analytics and big data. Data or Business analytics may simply be regarded as "*the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insights about their operations, and make better, fact-based decisions*" (Davenport and Harris 2007). Big data is any data whose volume, variety, velocity, and veracity is a challenge to process and understand (Cukier and Mayer-Schoenberger 2013). Both are being used increasingly by business clients, and these developments appear to be lasting. As such, the audit profession is beginning to examine how both developments affect the standards and practice. Public auditing in the modern economy may require a paradigm shift amd this dissertation provides several needed ideas for such a change.

These recent concerns have been recognized as follows and are identified here. Namely:

1. What does previous research say about analytics in the audit engagement?
2. Should new (modern) analytics methods be used in the audit process?
3. Which of these methods are the most promising?
4. Where in the audit are these applicable?
5. Should auditing standards be changed to allow / facilitate these methods?

6. Should the auditor report be more informative?

7. What are the competencies needed by auditors in this environment?

8. How can the provenance of external Big Data provide assurance as audit evidence?

These eight issues serve as the research questions for this paper and guide its organization. The background of practice and standards are first summarized in the second section of this Introduction, within the context of big data and analytics.

The second chapter concerns the first research issue, that of extant literature regarding Analytical Procedures (APs) in the audit engagement. However, as the second chapter notes, when examining the extant audit literature it appears that there are no studies that organize this research. Accordingly, extant research is organized in a process initially encompassing 572 papers that eventually ends with 301 papers regarding Analytical Procedures in the external audit.

The second chapter additionally addresses the following three concerns: should more complex analytics be used in the engagement? If so, where in the audit process are these most applicable? Which techniques appear to be most promising? This chapter proposes that the answers to these questions may be assisted by an examination of the extant external audit research. Accordingly, 301 papers are identified via the Systematic Literature Review Method (SLRM) that discuss the use of analytical procedures in the public audit engagement. These papers are categorized by technique, engagement phase, and other attributes for understanding. This analysis of the literature is constructed as an External Audit Analytics (EAA) framework, which is subsequently expanded with the concepts of business analytics (Holsapple et al, 2014). Specifically, this synthesis organizes the audit research, thereby offering guidelines regarding possible approaches for more complex and data driven analytics in the engagement.

The third chaper continues with the discussion of the next six research questions. Many potential directions for future research are suggested, based on the findings from the literature review and audit evidence prespectives.

The fourth chapter discusses the final research question that emerges, regarding how can the provenance of external big data sources may provide assurance as sufficient Audit Evidence. The standards regard external sources of evidence as being highly reliable. However, in this age of big data many sources of evidence are untraceable and their provenance unverifiable, such exogenous data may not be sufficient. This chapter proposes a solution for assuring the secure provenance of big data, allowing it to be regarded as reliable evidence for external auditors.

The fifth chapter concludes this discussion of external audting, analytics, and big data in the modern economy.

These chapters discuss and illuminate broadly many issues facing the profession since business clients are becoming increasingly automated and are capturing massive amounts of data. These chapters contribute towards solutions for several of these issues. These chapters also contribute to the stream of literature in the audit profession regarding external audit analytics and reliability of big data audit evidence. Hopefully these chapters encourage conversation and debate among academics, regulators, and the profession.

1.1 Background: Discussion of the Current External Audit Environment

> "Advances in technology and the massive proliferation of available information have created a new landscape for financial reporting. With investors now having access to a seemingly unlimited breadth and depth of information, the need has never been greater for the audit process to evolve by providing deeper and more relevant insights about an organization's financial condition and performance –while maintaining and continually improving audit quality.
>
> Does this mean that core elements of the audit such as the current "pass/fail opinion" that external auditors are mandated to provide – and that has served investors well for years, need to expand? Absolutely!" (Liddy 2014)[1]

There is an increasing recognition in the external audit profession that the emergence of big data (Vasarhelyi, Kogan, and Tuttle 2015) as well as growing use of data analytics in business processes has brought a set of new concerns and opportunities to the audit community. Accountants[2], Large Audit Firms[3], Standard

---

[1] James P. Liddy is KPMG LLP U.S. Vice Chair, Audit and Regional Head of Audit, Americas. Article published in Forbes August 4, 2014.

[2] The AICPA's Assurance Services Committee (ASEC) has met three times over the last three years with the Auditing Standards Board (ASB) to discuss audit analytics, and how the use of analytical tools and techniques fit within the current standards. As a result, the ASEC is developing a new Audit Data Analytics guide that will replace the current Analytical Procedures guide. The Audit Data Analytics guide will update and carry forward much of the content found in the Analytical Procedures guide, and will also include discussions around Audit Data Analytics and how they can fit within the current audit process. ASEC's Emerging Assurance Technologies task force is also working on a document that will map the traditional audit procedures to current analytical tools available today and elements of continuous audit.

[3] Every one of the "Big Four" has publicly announced efforts in the area of data analytics. Some have published white papers on the matter (e.g. Deloitte, "Adding insight to audit – Transforming Internal Audit through data analytics"; PwC, "The Internal Audit Analytics Conundrum—Finding your path through data"; KPMG, "Leveraging data analytics and continuous auditing processes for improved audit planning, effectiveness, and efficiency"; EY, "Big data and analytics in the audit process: mitigating risk and unlocking value").

Setters[4], and Academics [5] have been progressively raising many issues, among which we find:

1. What does previous research say about analytics in the audit engagement?
2. Should new (modern) analytics methods be used in the audit process?
3. Which of these methods are the most promising?
4. Where in the audit are these applicable?
5. Should auditing standards be changed to allow / facilitate these methods?
6. Should the auditor report be more informative?[6]
7. What are the competencies needed by auditors in this environment?
8. How can the provenance of external Big Data provide assurance as audit evidence?

These concerns have emerged even though analytical procedures in general have been addressed by the American Institute of Certified Public Accountants (AICPA) guidelines of 1972 and in numerous academic papers since 1955. The Statement on Auditing Standards (SAS) No. #1, states:

> "The evidential matter required by the third standard (of field work) is obtained through two general classes of auditing procedures: (a) tests of

[4] In April 2015, the IAASB started a subcommittee on analytic methods and heard presentations on the matter (e.g., Dohrer, Vasarhelyi, and McCollough 2015). The objectives of the subcommittee are to explore developments in audit data analytics and how the IAASB will respond to these developments. Also, the PCAOB has approached the "Big Four" to discuss the usage of analytics.

[5] A special section of Accounting Horizons with 7 articles (see Vasarhelyi, Kogan, and Tuttle 2015) has been dedicated to big data. An increasing number of articles in the accounting literature (see ensuing sections) have emerged proposing and illustrating analytic methods.

[6] The PCAOB issued Release No. 2016-003 on May 11, 2016 re-proposing new standards for the audit report in which in addition to the traditional pass/fail model "critical audit matters" (CAM) would be disclosed.

> details of transactions and balances, and (b) analytical review procedures applied to financial information (AICPA 1972 par. 320.70)."

There is a fine balance in every audit engagement between detailed evidence collection and analytical procedures (Yoon 2016). Detailed evidence collection can be quite costly yet deemed more reliable according to the standards, while analytical procedures are widely viewed as being less costly and believed less reliable by regulators (Daroca and Holder 1985; Tabor and Willis 1985). Both processes are allowed by the standards; their degree of utilization depends on auditor professional judgment. While the requirement of tests of details of transactions and balances is somewhat defined, the second requirement of analytical review procedures is completely undefined, except that it should be applied to financial data (Tabor and Willis 1985).

More recently, according to AU-C Section 520 about Analytical Procedures (AICPA 2012a), to conduct substantive analytical procedures the auditor should:

- determine the suitability of a certain substantive procedure, given the account;
- evaluate the reliability of the data from which these ratios are developed;
- develop an expectation of recorded amounts and ratios and whether these are accurate, and finally
- determine the amount of difference (if any) between the recorded amounts and the auditor's expected values and
- decide if the difference is significant or not.

The lack of detailed recommendations in this age of automation and big data regarding which analytical procedures to undertake in the external audit engagement has inspired considerable discussion. Although the internal audit environment is increasingly using analytics (Vasarhelyi et al. 2015; Perols and Lougee 2011; Dilla et al. 2010; Yue et al. 2007; Alles et al. 2006; Church et al. 2001), the external audit field has not responded to the same degree. The regulations, such as the guidance for sampling, have remained unchanged despite the fact that many audit clients automate the collection and analysis of 100% of their transactions (Schneider et al. 2015; Zhang et al. 2015).

1.2 Background: Current Practice and the Standards

It is essential to understand the current scope and constraints of the public audit profession before envisioning the role of more complex analytics and big data in the engagement. Since auditing is largely a regulation driven profession, the expectations regarding evidence collection and analytical procedures should be considered. The auditor still needs to test for basic assertions to make sure that the objectives of the audit are fulfilled regardless of the nature of the evidence and the way the evidence is being collected. The tests for certain assertions may change in the current new environment with its different nature of evidence and the way this evidence is collected and analyzed. However, even if the tests of assertions were to be altered, the assertions themselves wouldn't change and neither would the fundamental objective of the public auditor – to provide opinion on the financial statements as to whether they represent the financial position of the client in accordance with the generally accepted accounting principles.

1.2.1 Analytical Procedures and the Standards

Analytical procedures are required by the Public Company Accounting Oversight Board (PCAOB) in the planning phase (PCAOB 2010, AS No. #2110) and review phase (PCAOB 2010, AS No. #2810), but are undertaken according to auditor judgement in the substantive procedures phase (PCAOB 2010, AS No. #2305). The PCAOB asserts that analytical procedures can range from scanning, simple comparisons, and ratio analysis to more complex models involving many types of data elements and their relationships. Furthermore, the PCAOB states in AS No. #2305.03: "An understanding of the purposes of analytical procedures and the limitations of those procedures is also important."

The purpose of analytical procedures is different for each audit phase. For the risk assessment/planning phase, analytical procedures should enhance the auditor's understanding of the client's business and its transactions or events, and identify areas that may indicate particular risks to the audit. The auditor is expected to perform analytical procedures for the revenue accounts, to reveal unusual relationships indicative of possible material misstatements. The auditor should also use his or her knowledge of the client and its industry to develop expectations. The standards admit that the data may be at a more aggregated level and result in a less precise analytical procedure which is still acceptable at this phase. It would appear that the standards do not preclude the use of exploratory or confirmatory analytics in this phase, whether simple or more complex.

According to AS No. #2305.04, analytical procedures are used in the substantive testing phase to obtain evidence about certain assertions related to certain accounts or business cycles. Analytical procedures may be more effective than tests of details in some circumstances (Yoon 2016). In AS No. #2305.09, the PCAOB states that "the decision about which procedure or procedures to use to achieve a particular audit objective is based on the auditor's judgement on the expected effectiveness and efficiency of the available procedures." The main limitations appear to be the "availability" of certain procedures and the auditor's judgement on the expected effectiveness of certain analytical methods. The latter condition would appear to reflect the auditor's level of familiarity with certain analytical methods.

For the review phase of the audit engagement, analytical procedures are required to evaluate the auditor's conclusions regarding significant accounts and to assist in the formation of the audit opinion (PCAOB 2010, AS No. #2810.05-.10). Similar to the planning phase, the auditor is required to perform analytical procedures related to revenue during the relevant period. In this section, there is no mention of particular analytical approaches, except that this phase typically is similar to the planning phase. As such, it is expected that the more complex exploratory or confirmatory techniques are not excluded here either (Liu 2014).

1.2.2 Evidence Collection and the Standards

The main purpose of the work conducted by an auditor in an external engagement is to obtain reasonable assurance that the client's financial statements are free from material misstatements and to subsequently express an opinion regarding these financial statements and the client's internal controls in the auditor's report. To accomplish this task, the auditor must design and perform audit procedures to obtain sufficient appropriate evidence; furthermore, the Audit Standards require auditors to examine physical evidence as part of the risk assessment process (PCAOB 2010, AS 1105; AICPA 2012, SAS 122; IAASB 2009, ISA 500). Audit evidence is all the information (whether obtained from audit procedures or other sources) that either confirms or contradicts or is neutral about management's assertions on the financial statements or internal controls.

Additionally, the Sarbanes-Oxley Act (SOX) demands that public accounting firms maintain the records of an audit report (and all of its supporting information) for at least seven years after its issuance (United States Public Law No. 107-204; Tackert et al. 2004). The Sarbanes-Oxley Act also mandates that auditors verify the accuracy of the information or evidence that forms the basis of their audit opinion. Since SOX, audit firms have relied more heavily on detailed audit examination, ratio analysis, and scanning for substantive analytical procedures as these are regarded to be "harder" audit evidence formats than regression and other "softer" analytical techniques (Glover et al 2014). The impact of this legislation on the profession's analytical procedures choices should not be ignored. However, as mentioned and footnoted in the Introduction, every one of the "Big Four" has recently publicly announced efforts in the area of data analytics for assurance services.

Since audit evidence is all the information used by the auditors to form the audit opinion (PCAOB, 2010, AS 1105), it should be both sufficient and appropriate. Sufficiency is the measure of the quantity, the amount of which is determined by detection risk determined by the auditor and the level of quality of the evidence, or it's appropriateness (PCAOB 2010, AS 1105). Appropriateness is the measure of relevance (what does the evidence tell the auditor) and reliability (can the auditor trust the evidence)? Basically, if the underlying information is not reliable and its origin isn't verifiable, then more evidence will need to be collected and reviewed (Appelbaum, 2016). Poor quality evidence cannot be compensated for by collecting a larger amount of data (PCAOB 2010, AS 1105).

However, in today's complex IT and big data environment, the nature and competence of this audit evidence has changed (Brown-Liburd and Vasarhelyi 2015; Warren et al. 2015; Nearon 2005). With big data, quantity of evidence is hardly an issue with which to be concerned. However, quality of electronic evidence becomes even more dominant in the equation and may be more challenging to verify. Most stages of a transaction can be computer generated and recorded and can only be verified electronically. For example, with additional information available from external big data, intangible assets might be partially valued by the client from information derived from text analysis of aggregated tweets and web scraping of social media. However, the reliability of these tweets and social media is hard to verify (Appelbaum 2016).

The issues for electronic accounting data and electronic audit evidence are drastically different from that of manual and paper-based examination. Many of the characteristics that are strengths with paper-based evidence pose issues for electronic evidence. Where paper documentation is regarded as not easily altered, electronic data may be easily changed and these alterations might not be detected, absent the appropriate controls. In paper-based evidence collection, sources that are verified external to the client are considered to be highly reliable (PCAOB 2010, AS 1105), whereas external electronic evidence is difficult to verify for origin and reliability. Paper-based evidence is easy to evaluate and understand, whereas electronic data and evidence may require a high level of technical expertise of the auditor. Since big data is electronic data, big data presents a scenario where these characteristics are magnified by many degrees. Furthermore, the types of tests that should be undertaken by auditors to examine basic assertions may change.

Auditors are required to conduct the audit engagement within the parameters of the regulations, regardless of the IT or accounting complexity of the client. It is highly probable that the client may be undergoing processes with advanced analytical techniques and new sources of data. The newest challenges facing the auditor are the increasing use of big data and the subsequent application of more advanced analytics by clients.

After gaining an understanding of this current audit environment of big data and advanced analytics, what follows are immediate research questions that should be addressed if the profession is to integrate itself within this new business paradigm.

CHAPTER TWO: DATA ANALYTICS FOR EXTERNAL AUDITING: A COMPREHENSIVE LITERATURE SURVEY

2.0 Introduction

There is increasing recognition in the public audit profession that the emergence of big data as well as the growing use of analytics by audit clients has brought new opportunities and concerns. That is, should more analytics be used in the engagement and if so, where (Issues 3,5#)? Which techniques appear to be most promising (Issue 4#)? More importantly:

ISSUE 1: What has been the research to date regarding the use of analytics in the audit engagement ?

Before these many issues can be addressed, researchers should understand the scope of extant research.

The standards do not explicitly define the type of analytical approaches that should be undertaken by auditors to fulfill regulatory requirements, except that the auditor should develop an expectation from the appropriate analytics of reliable data from certain accounts, and then calculate the difference of these expectations and the recorded numbers (AS 2305, PCAOB, 2016). The standards require that analytical procedures be undertaken in addition to evidence collection at the preliminary review and final review stages (Daroca & Holder, 1985), but the decision about which analytical approach techniques to use are left to auditor judgment.

The opaqueness of this aspect of public auditing has led to numerous debates and discussion within the auditing academic community since 1958 (AICPA 1958). These debates have increased with the emergence of big data and automation of business financial reporting (Vasarhelyi, Kogan, and Tuttle 2015). These discussions and debates, as evidenced in academic publications, are indicative of the degree and breadth of analytical approaches available to the engagement. Therefore, it is only natural to investigate this vast body of audit research for insights regarding an expanded use of analytics. This research is relevant to:

- ✓ Audit academics and researchers who are interested in continuing with new research about analytics in the external audit engagement and who can refer to this paper for guidance as to which areas have previously been discussed in the literature and which could benefit from additional attention
- ✓ Practitioners or auditors who want to be aware of the degree of research and of innovative ideas about analytics and to possibly incorporate them in the engagement
- ✓ Regulators who are seeking to update the standards and suggest best practices regarding the use of analytical procedures in the audit engagement.

This chapter is an attempt to identify and categorize publications referencing the use of analytics in the engagement. Accordingly, 301 papers are evemtually identified that discuss some aspect of analytical procedures in the external audit engagement. The large number of papers make it difficult for academics and practitioners to identify specific analytic techniques or gaps in the research. Therefore, these papers are then categorized by technique, engagement phase, and other attributes to facilitate an understanding. This analysis of the literature is subsequently categorized as an External Audit Analytics (EAA) framework, the objective of which is to identify gaps, to provide motivation for new research, and to classify and outline the main topics addressed in this literature. Specifically, this synthesis organizes audit research, thereby offering guidelines regarding possible future research into more complex and data driven analytics.

Following this Introduction, the Background section discusses Analytical Procedures as promulgated by the standards and purportedly practiced by the profession, in contrast to the complex Business Analytics that are being progressively utilized by engagement clients. The third section begins the Literature Review process by discussing the methodology for collecting these papers and how they are categorized by timeline, research methods, audit stage, technique, and orientation. The fourth section discusses the meaning of the results of the literature review, areas for future research, and gaps in the literature. An External Audit Analytics (EAA) conceptual framework is proposed to facilitate an understanding of not only where research has been undertaken but also, given an understanding of business analytics practices by audit clients, where future research should concentrate. This visionary EAA conceptual framework is derived from the synthesis of the literature in the context of business analytics. This chapter then concludes with implications and discussions for future research regarding the broad potential for analytics in the external audit.

2.1 Background

2.1.1 Analytical Procedures in the Standards and Typical Practice

AS 2305 (PCAOB 2016) defines Analytical Procedures (APs) as an "important part of the audit process that consists of evaluations of financial information made by a study of plausible relationships among both financial and nonfinancial data." AS 2305 states that APs may range from basic comparisons to the use of more complex models involving multiple relationships and elements in the data. APs are required in the planning/risk assessment phase and in the review phase of the engagement. APs utilized in the preliminary planning/risk assessment phase are typically considered as reasonableness tests. At the review stage of the audit, they provide an overall review of the assessments and conclusions reached. APs may be used as a substantive test to obtain evidence about certain assertions related to account balances or types of transactions. In certain circumstances, APs may be more effective and efficient than substantive tests of details. When the data set is large and varied, APs may be more effective. When the risk of misstatement is minimal, APs may be more efficient and less costly.

The Cushing and Loebbecke (C-L) model (Figure 1) reflects the phase structure of the typical audit engagement by the Big 8 firms at that time and is the basis for the audit model in many textbooks (Louwers et al. 2016; Whittington and Pany 2014). In this model, auditors should conduct a preliminary analytical review in the planning activities, conduct analytical review procedures as well as substantive tests of transactions and tests of balances in the substantive testing phase. In the evaluation and review phases, this work requires revisiting and re-performing analytical tests (Cushing and Loebbecke 1986). Continuous Activities seemed to consist primarily of project management duties, light documentation, and follow-up procedures.

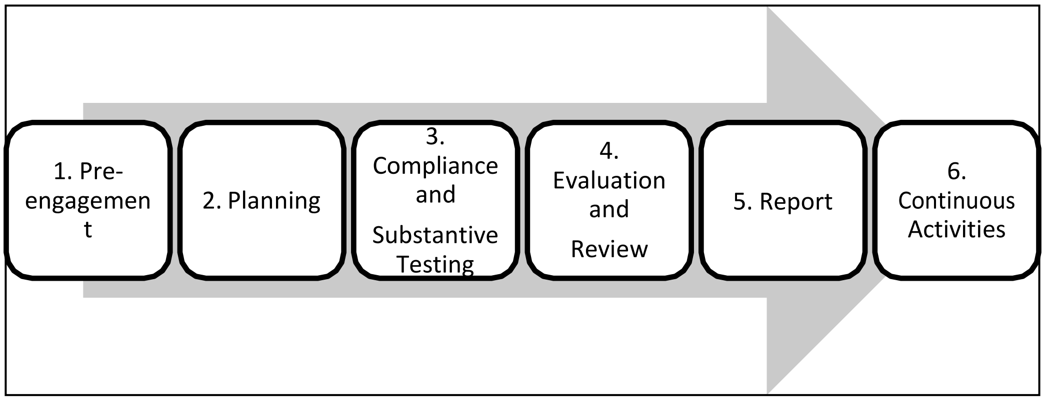

Figure 1: The Six Stages of the Audit Cycle (Cushing and Loebbecke 1986)

In the substantive phase, where the auditor examines the data with some combination of APs and tests of details, the auditor may use a sampling approach to select transactions to examine. According to AS 2315 (PCAOB 2016), "Audit sampling is the application of an audit procedure to less than 100 percent of the items within an account balance or class of transactions for the purpose of evaluating some characteristic of the balance or class." The sampling process requires much consideration, judgment, and planning by the auditor. The sampling procedure may be statistical or non-statistical.

As described in AS 2305.05 (PCAOB 2016), analytical procedures "involve comparisons of recorded amounts, or ratios developed from recorded amounts to expectations developed by the auditor." For example, APs typically accomplish the following five tasks (Table 1):

Analytical Procedures	Sources of Information
Comparison of current year account balances to same account balances of other periods	Financial account information/reports
Comparison of current account balances to the anticipated results found in the client's budgets and forecasts	Client budgets and forecasts
Evaluation of the relationships of current year account balances to other current year balances for conformity with predictable patterns based on the client's experience	Financial relationships among accounts in the current period
Comparison of current year account balances and financial relationships (ratios) with similar information for the client's industry	Industry statistics
Study of the relationships of current year account balances with relevant nonfinancial information	Pertinent nonfinancial information

Table 1: Typical AP Engagement Tasks, adopted from Louwers et al (2015) pg 99

Based on this understanding of APs, now the literature may be reviewed for relevant papers and organized for ease of understanding. However, as will be discussed in the following section, the literature about APs is not confined to the fundamental processes described in Table 1, but instead is much broader and varied in scope, thereby complicating this task. This complexity requires an established system for organization, such as the Systematic Literature Review Research Method (SLRRM).

2.2 Literature Review

2.2.1 Systematic Literature Review Research Method

Keele (2007 p 3) states that "A systematic literature review…is a means of identifying, evaluating, and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest." Systematic research is conducted to:

- ✓ Summarize and organize the existing research
- ✓ Identify gaps in this research
- ✓ Provide a framework/background to understand the research and to appropriately direct new research activities

A systematic review synthesizes the research in a pre-disclosed search and organization strategy that is auditable and unbiased. The systematic review process begins with a discussion of the strategy that guides the research. This defined search strategy aims to detect as much of the relevant literature as possible. Keele (2007) suggests that the research protocol include:

- ✓ The research questions and topics that this study aims to address
- ✓ Methods, sources, and techniques used in the identification of relevant papers such as key words, search strings, digital search engines, libraries, journals, and conferences
- ✓ Inclusion and exclusion criteria
- ✓ Attribute assessment process for the extracted literature
- ✓ Procedures necessary to develop a research-directing framework

The remainder of this sub-section about the systematic literature review presents these listed protocol features.

2.2.1.1 Objectives and Research Questions

The main objective of this research is to explore and then categorize and synthesize the research on analytical procedures in the external audit engagement. In this context, the primary concern of the profession is whether business analytics should be used in the engagement, and if so, when and how often? And should these techniques be more complex? However, it is not yet ascertained that these are concerns of academics historically. Accordingly, the first research question is:

> RQ1: What are the main research topics and aspects covered by the research about analytical procedures in the external audit engagement?

Then, building on the recent concerns of the profession and the information from RQ1, these research questions ensue:

> RQ2: How do researchers propose that analytical procedures be applied in the external audit engagement?
>
> > RQ2.1: What is the time line for this general research topic?
> >
> > RQ2.2: Which research methods are being utilized more frequently by academics?
> >
> > RQ2.3: How many papers have been published about analytical procedures in the external audit engagement?

RQ2.4: In which journals have these papers been published?

RQ2.5: When in the audit engagement do researchers propose that analytical procedures be applied?

RQ2.6: How often do researchers propose that analytical procedures be applied?

RQ2.7: What type of analytical procedures do they suggest be used?

The third objective is to organize these selected papers in a structured framework which can assist in organizing this literature and identify existing gaps and areas for further investigation.

RQ3: How to organize the main attributes covered by these studies of analytical procedures in the audit engagement?

The fourth objective is to organize the literature in a structured framework that can appropriately direct future research activities:

RQ4: Given the attributes categorized in RQ3, how can this literature be presented to direct future research?

2.2.1.2 Search Strategies

Having determined the general research questions, the search strategies, search parameters, and search sources can now be defined.

Keywords: Keywords and search strings are collected based on the research questions. This process entailed keyword searches for “analytics”, “analytical procedures”, "analytical review", "audit planning", "risk assessment", "internal control assessment", "compliance testing", "statistical analysis", "statistical sampling", "substantive testing activities", "review", "fraud", "Going Concern", and "Fair Value Assessment". Every technique type was also included in the search, as listed in Table 21.

Search strings: These are constructed from the keywords in conjunction with the research questions. The string format is generic so that it may be used in most libraries. For example: (Management Fraud) OR (Earnings Misstatement).

Sources: To accomplish the task of initially identifying relevant papers, the database of auditing research compiled by a sub-committee of the AAA Auditing Section Research Committee (Trotman et al, 2009) is examined for academic papers likely to discuss audit analytics. The references of these papers are also examined for likely additions to the list and those subsequent papers are similarly reviewed and additional references tracked, in an iterative process. This entire process is then repeated in Google Scholar and SSRN.

2.2.1.3 Selection Criteria

The papers selected for this study had to be published as full papers in academic journals or as completed dissertations or as completed working papers published online. After obtaining the results from the inclusion/exclusion lists that follow, all remaining studies were examined again for the required additional textual analysis. Table 2 shows the selection steps for the literature review.

Selection Step:	
Step 1	Apply keywords and strings to all sources and follow up with source references, gathering results until additional papers cannot be extracted
Step 2	Exclude any invalid papers
Step 3	Apply inclusion/exclusion criteria to titles, keywords, and abstracts
Step 4	Apply criteria to introductions and conclusions
Step 5	Review the entire text, applying exclusion/inclusion criteria

Table 2: Format of literature selection process (Keele 2007)

The complete table of all identified papers and major categorizations can be found in Appendix B. The inclusion criteria are as follows:

- ✓ Papers published in academic journals, completed dissertations available online, and working papers published online
- ✓ Papers mentioning external auditing, audit engagement, assurance services, engagement team, public accounting/auditing, financial auditing
- ✓ Papers discussing some aspect of analytical procedures/analytics/statistics/sampling/data mining/machine learning and/or one of those techniques
- ✓ Papers discussing at least one phase of the audit (see discussion that follows)
- ✓ Papers where analytics are not the primary focus but meet all other criteria (this is typical for many behavioral studies)

Papers are excluded based on the following criteria:

- ✓ Papers published in media that were practitioner journals at the time of publication

- ✓ Conference papers and workshop papers
- ✓ Incomplete papers and duplicate papers
- ✓ Papers that mention "auditing" or "auditor" but do not distinguish internal from external and do not describe or refer to a typical engagement responsibility or task
- ✓ Papers referring only to internal auditing/auditors
- ✓ Papers that do not discuss some aspect of analytics/statistics/sampling/data mining/machine learning and/or one of those techniques as either primary or secondary focus
- ✓ Papers that discuss some aspect of a technique but don't relate it at all to auditing (for example, papers on MU sampling never mention auditing or an audit phase or function)

In general, a paper is considered relevant if it mentions directly external auditing and discusses an aspect of analytics that typically belongs in at least one phase of the external audit model as developed by Cushing and Loebbecke (1986), see Figure 1 (Elliott, 1983). In the public company audit setting, analytics could be the primary focus of the paper or a secondary focus or part of another process/objective. For those papers where the use of analytics is not the primary focus, only those papers where analytics are essential to the process/argument/study are selected. For example, several behavioral studies are included that focus on professional judgement and utilize analytical procedures in the experiment or survey process (e.g. Arrington et al, 1984; Asare and Wright 1997). Furthermore, if an analytical procedure is discussed but the typical stage of the audit cycle for that procedure is not identified directly by the author(s) but is otherwise described, the audit cycle is not identified in the categorization table in Appendix B. For example, there are several early papers which extensively discuss statistical sampling and substantive testing but never mention the substantive procedures phase, so this stage of the audit cycle is not listed with those papers in the categorization table in Appendix B.

This literature selection process encompasses a total of 572 papers across auditing, systems, accounting, economics, and finance literature and after applying the selection process, results in 301 papers. The entire texts of the excluded 271 papers were then examined to determine that they truly do not qualify (Table 3).

Exclusion Reason	**Number of Publications Excluded**	**Running Total Number of Included Publications**
Total Number of Papers		572
No mention of EXTERNAL or PUBLIC Audit/phase	(103)	469
Not available online (usually these are references from earlier publications)	(47)	422
APs are not mentioned	(21)	401
All other exclusion reasons	(100)	301
Total Exclusions	**(271)**	**301 (Total of Inclusions)**

Table 3: Reasons for Literature Reduction

2.2.2 Literature Categorizations addressing the Research Questions

2.2.2.1 (RQ1)What are the main research topics and aspects covered by the research about analytics in the external audit enagegement?

A large majority of the papers (80%) discuss the effectiveness or efficiency of various APs as the primary topic. Fourteen papers mention the effectiveness and efficiency of the APs as topics for future research. The overwhelming thrust of each paper is the quality of the performance of APs as either a primary or secondary factor in some aspect of the external audit (Table 4).

In summary, RQ1 determines that the performance of APs in the audit engagement is the predominant concern of this body of literature and it supports the main objective of this research paper.

Focus of Research	**Number of Papers**
AP use in different phases, internal controls, sampling, and evidence	177
AP as secondary emphasis to primary topics such as judgment, independence, bias, and experience	60
APs to detect earnings misstatements and management fraud	28
Fraud detection (employee and financial statement)	14
Going Concern/Bankruptcy Assessments	18
APs for Valuations	4

Table 4: Research Focus of the papers

2.2.2.2 (RQ2) How do researchers propose that analytical procedures be applied in the external audit engagement?

This research question relates to the main objective of this study, which is to determine how extant audit research applies analytical procedures to the engagement. That is, should more analytics be used in the engagement and if so, where? What techniques appear to be the most promising? The many angles of this query are addressed in the sub-research questions that follow.

2.2.2.3 (RQ2.1) What is the Timeline for this research?

Most research about APs in the financial audit engagement appears to be accessible online for publications as of 1958. Although the publications were sparse for the first two decades, this changes in the 1980's and maintains that pace ever since for a total of 301 papers (Figure 2).

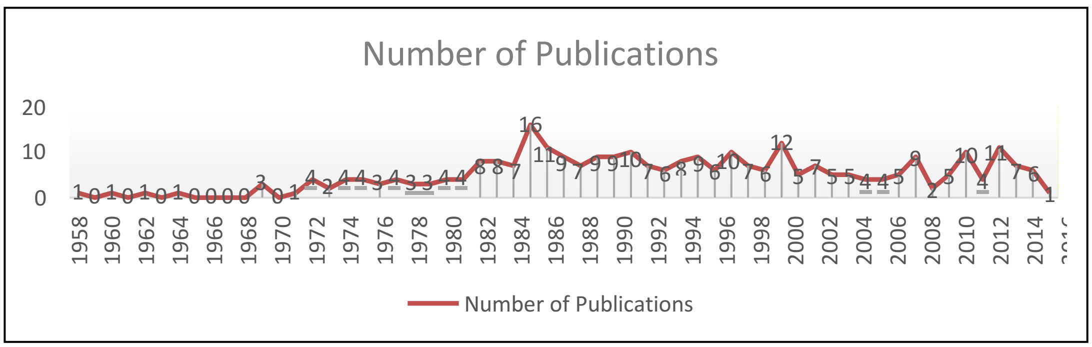

Figure 2: Detailed timeline of academic literature discussing AA in the external audit environment

2.2.2.4 (RQ2.2) What are the Research Methods of the Literature?

These papers are also classified by their research method into the following categories:

- ✓ Analytical (Case Study, Design Science, Empirical)
- ✓ Behavioral (Education Case Study, Experimental, Field Study, Survey)
- ✓ Archival (Literature Review, Historical)
- ✓ Conceptual (Discussion, Theoretical, Normative)

The research methods are described more precisely per paper in Appendix B, but are summarized in the body of this manuscript at the level of Analytical, Behavioral, Archival, and Conceptual, since these general approaches are predominant. For example, a paper may be classified as a survey in Appendix B but be represented in this figure as behavioral. A summary of these classifications is reported here in

Figure 3. The Analytical, Behavioral, and Conceptual approaches are equally popular, with the Archival approach being undertaken minimally in comparison.

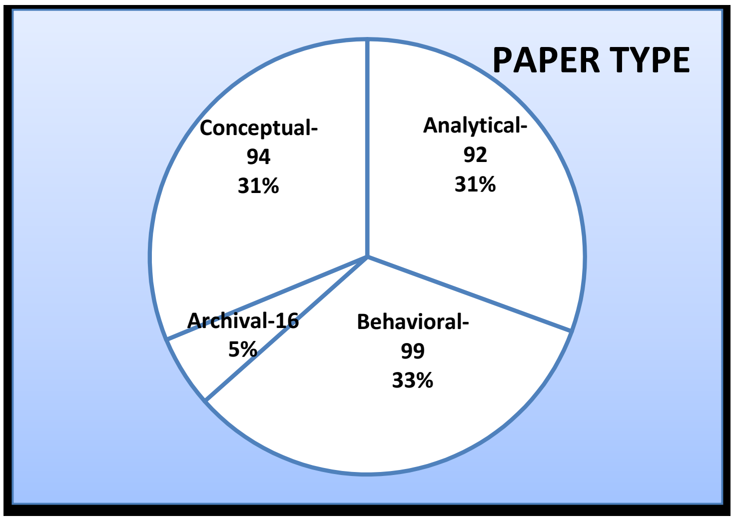

Figure 3: Display of the number of paper types/approaches that discuss analytics in the external audit

These 301 papers vary in both research methods and in analytical techniques. The most popular research methods are analytical, behavioral, archival, and conceptual. The use of these four research methods is compared below in Figure 4, with a more detailed comparison and separate analyses in Figure 27, Figure 28, Figure 29, Figure 30, and Figure 31 of Appendix A.

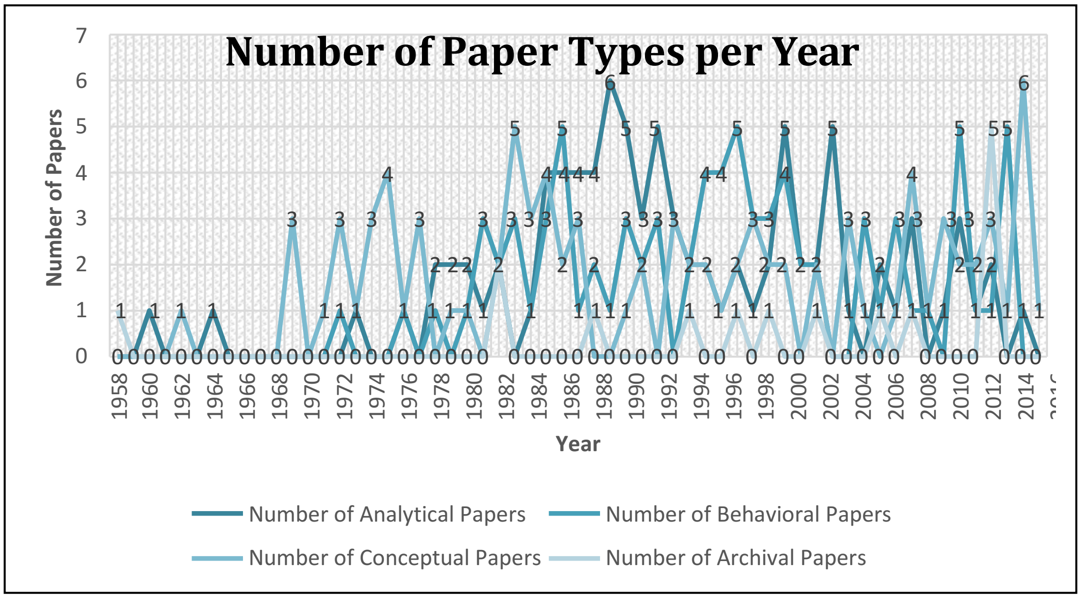

Figure 4: Comparison of the four main research methods in external audit analytics

2.2.2.5 (RQ2.3) What is the Background Story of Analytical Procedures Research and how many papers were published?

In November 1956, the American Institute of Certified Professional Accountants (AICPA) formed a special AICPA committee, the Committee on Statistical Sampling (AICPACSS), reflecting the growing dissatisfaction of auditors with the use of judgement based sampling processes (AICPA, 1958). Although the American Institute of Accountants (preceding the AICPA) published a 134 page book in 1955 that detailed eight sampling application case studies, titled *A Case Study of the Extent of Audit Samples* and which was also discussed in Elder et al (2013), Weber (1978), and Joyce (1976), these case studies merely highlighted the differences in sampling approaches. There was a perceived need for a more objective and scientific approach for deciding the number of items to be tested when performing audit procedures (Tucker and Lordi, 1997).

Although it cited over 118 papers on sampling that had been published during the previous ten years (e.g.. Neter 1949, Arkin 1957, Arkin 1958, and Hill 1958), the 1958 AICPA review that emerged from this committee is regarded as the first position paper regarding consideration of statistical sampling in audit procedures (Tucker and Lordi, 1997). The Committee's study revealed that until the mid-1950's, there was scant knowledge among auditors regarding statistical sampling, even though it was being used with greater frequency by large public accounting firms (AICPA, 1958). Auditors were using primarily "block testing", where a period of time was selected and audit tests applied to that period, or judgment based sampling, where the sample was extracted based on client/industry/professional expertise.

The AICPA study could probably be considered the first that discusses more advanced APs to replace simple calculations and judgement based external auditing procedures. After this 1958 publication, there were several years of sporadic publications regarding APs in the external audit environment (Figure 5). During the 70's there were several years with peaks of four publications, followed by a flourish of activity in the 80's and 90's. It is during this period that the use of ratio analysis was challenged, given its relative problematic accuracy (Deakin 1976). The research regarding external audit analytics peaked in 1985, at 16 publications. Another paper questioning the over reliance by the profession on relatively ineffectual ratio analysis was published (Glover, Prawitt and Wilks 2005). The years 2000, 2011, and 2013 experienced smaller peaks in activity. A more detailed graph (*Figure 27*) showing publications per year can be found in Appendix A.

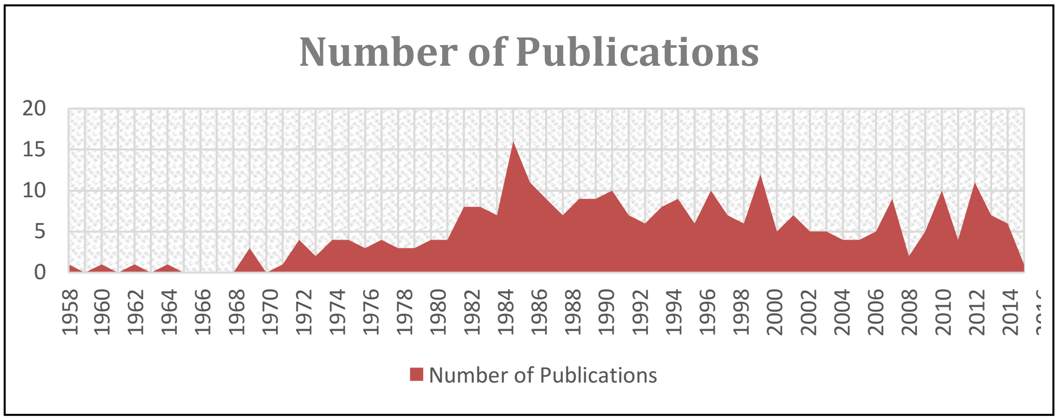

Figure 5: Number of papers published per year that discuss analytics in the external audit setting

2.2.2.6 (RQ2.4) In which Journals have these papers been published?

The papers are published in thirty-three different journals, with *Auditing: A Journal of Practice and Theory* with the higher frequency, followed by the *Accounting Review,* the *Journal of Accounting Research,* and *Contemporary Accounting Research*. Figure 32 in Appendix A displays the number of papers published by each journal. The earliest papers were published primarily in *The Journal of Accountancy* and *The Accounting Review*, both of which were considered to be the primary academic accounting publication venues at that time (Vasarhelyi, 1982). Prior to and changing in the 1950's, accounting academic literature emphasized individual expert opinion (most papers were single authorship) and internal logic (Vasarhelyi 1982; Vasarhelyi et al, 1988). Academic accounting research evolved during the late 50's and early 60's into more empirical thought and interdisciplinary approaches (Vasarhelyi 1982). Prior to the advent of the *Auditing: A Journal of Theory and Practice,* many papers referred to auditors as "outside accountants" or as "accountants and auditors" (Keenoy, 1958; Arkin, 1958; Hill, 1958). Auditing became more established as a field of its own, with unique issues of judgment and expertise that frequently were examined with behavioral methods (Felix and Kinney, 1982).

Specific areas of emphasis for analytical review procedures in the external audit are shown in this literature to be Financial Statement/Management Fraud (Hogan, Rezaee, Riley Jr & Velury, 2008; Trompeter, Carpenter, Desai, Jones & Riley Jr, 2012), Going Concern Opinion (Carson, Fargher, Geiger, Lennox, Raghunandan & Willekens, M., 2012), and Fair Value Measurement (Martin, Rich, & Wilks, 2006; Bratten, Gaynor, McDaniel, Montague & Sierra, 2013).

Additionally, statistical sampling is mentioned in 164 of the papers, which could be expected given the importance of this topic in the application of analytics during the last 50+ years. Analytical issues in sampling motivated the AICPA to form its first commissioned committee in 1956. Figure 6 portrays a trend analysis of statistical sampling. Over time, statistical sampling research peaked in the early 80's and again around 2000.

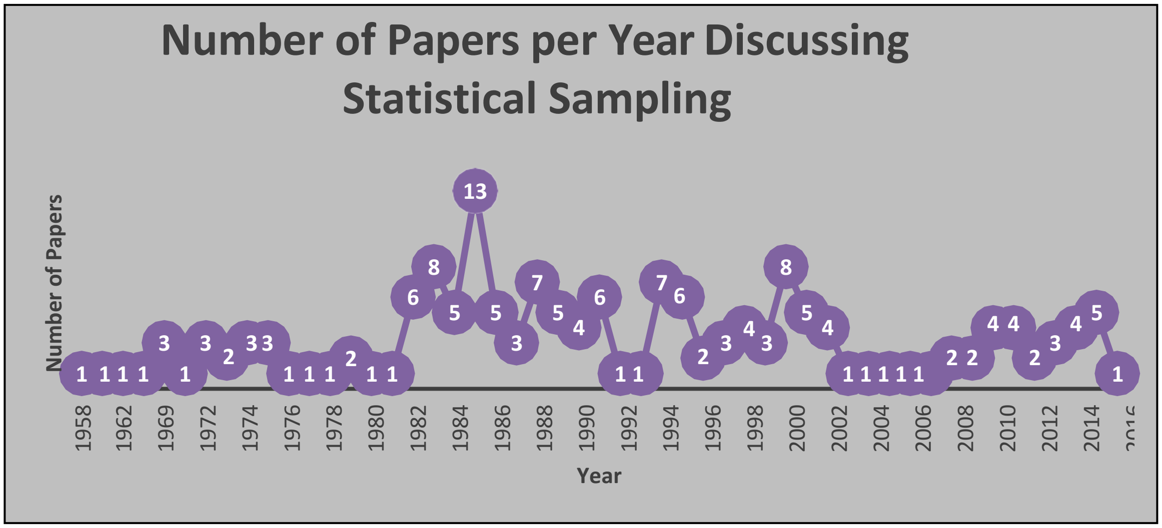

Figure 6: Analysis of the number of papers per year that discuss statistical sampling based methods.

2.2.2.7 (RQ2.5&6) When and how often should analytical procedures or analytics be applied?

The papers mention analytical methods in the six audit phases with the frequency shown below in Figure 7. Many papers discuss applying analytical methods in more than one phase, and each phase is separately counted. Analytics are discussed in the papers as follows: 36 times for the Engagement phase, 228 times for the Planning/Risk Assessment Phase, 225 times for the Substantive Testing Phase, 167 times for the Review Phase, 46 times for the Reporting Phase, and not at all in the Continuous Activities Phase. Given the role of analytical procedures as prescribed in the standards, it is not surprising that research is primarily concentrated in the phases of planning, substantive testing, and review and minimally in the areas of engagement and reporting.

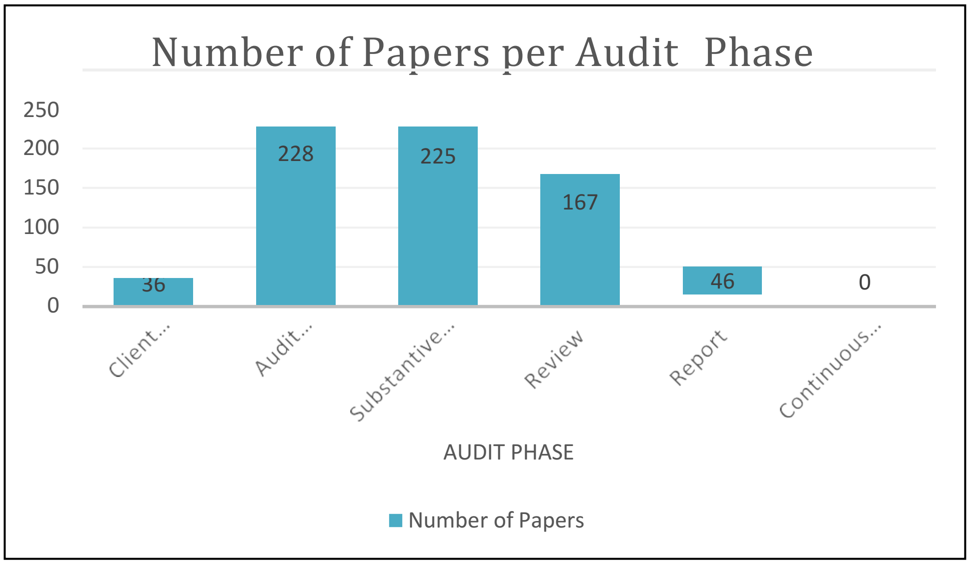

Figure 7: Number of Papers discussing the application of analytics per Audit Phase

2.2.2.8 (RQ2.7) What type of Analytical Procedures do they suggest be utilized by auditors?

The analytical procedures are also examined for each step of the C-L model (Figure 1). All phases are found to be similar regarding the inclusion of the Audit Examination techniques, as these procedures typically serve as a foundation for the application of more complex techniques.

The Audit Examination, Unsupervised, Supervised, Regression, and Other Statistical techniques are considered appropriate if they had been applied in the context of the Cushing-Loebbecke model (Figure 1), which may also be referred to as the "traditional" external audit model. A complete listing of the literature with audit phases and analytical techniques identified may be found in Appendix B. Furthermore, where papers mention audit assertions and auditor characteristics, these attributes are categorized. Other categories of classification include audit objectives, details of risk assessment procedures, details of substantive testing, details of internal control evaluation, resulting research questions, data quality and reliability, keywords, abstract/summary and results/conclusion as established and discussed by the papers. Every attempt is made to categorize these attributes exactly as they appear (or not) in these papers, without interpretation or inference of information.

Many of the techniques are applied to the different phases of the external audit, albeit sporadically in the case of unsupervised and supervised methods and frequently in the case of Audit Examination techniques and Regression techniques. Each of the audit phases of Engagement, Planning/Risk Assessment, Substantive & Compliance Testing, Review, Opinion Formulation and Reporting, and Continuous Activities exhibits academic research as follows (please see Table 21 in Appendix A and Appendix B for more detailed analysis per publication):

1. Engagement: The papers from this phase primarily discuss ratio analysis, regression, descriptive statistics, and expert systems, with only a few papers handling visualization, text mining, expert systems, multi-criteria decision aids and structural models.
2. Planning/Risk Assessment: Most of the papers in this phase deal with all types of audit examination, all of the regression techniques, and descriptive statistics, with some discussion of expert systems, Bayesian Belief Networks (BBN), and probability models, and slightly less of clustering, text mining, visualization, multi-criteria decision aids, and structural models.
3. Substantive Testing & Compliance Testing: Audit examination techniques are enormously popular here as were all of the regression techniques, descriptive

statistics, expert systems, BBN, and probability models. Less popular were all of the unsupervised method[1] and other supervised techniques[2]such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), genetic algorithms, bagging/boosting, and multi-criteria decision aids.

4. Review: Ratio analysis and Computer Assisted Audit Techniques (CAATS) are discussed frequently as were linear and time series regression and expert systems, with BBN, probability models, and descriptive statistics used occasionally.
5. Opinion Formulation and Reporting: In the opinion phase, the main techniques mentioned are ratio analysis, visualization, expert systems, log and linear regression, descriptive statistics and multi-criteria decision aids.
6. Continuous Activities: None of the papers discuss analytics in the context of ongoing/continuous activities.

All the techniques observed even once in the literature are marked in Table 5 below as to which audit phase they occur. Table 21 in Appendix A contains a listing of the papers for each technique per audit phase that were identified in the external audit literature.

Techniques:	**Audit Examination**	**Unsupervised**	**Supervised**	**Regression**	**Other Statistics**
Audit Phase:					
Engagement:	Ratio Analysis	Visualizations	Expert Systems/ Decision Aids	Log Regression	Multi-criteria Decision Aid

[1]Unsupervised approaches are those techniques that draw inferences from unlabeled datasets in which instances either have no output specified or the value of the output is unknown (such as whether a transaction is fraudulent or not)

[2] Supervised approaches are those techniques that draw inferences from labeled datasets, otherwise known as training data

		Text Mining		Linear Regression	Structural Models
				Time Series	Descriptive Statistics
				Univariate and Multivariate	
Planning:	Transaction Tests	Clustering	Process Optimization	Log Regression	Multi-criteria Decision Aid
	Ratio Analysis	Text Mining	Expert Systems/ Decision Aids	Linear Regression	Descriptive Statistics
	CAATS	Visualizations	BBN	Time Series	Structural Models
			Probability Model	ARIMA	
				Univariate and Multivariate	
Substantive & Compliance Testing:	Transaction Tests	Clustering	Process Optimization	Log Regression	Multi-criteria Decision Aid
	Ratio Analysis	Visualizations	SVM	Linear Regression	Benford's Law
	Sampling	Text mining	ANN	Time Series	Descriptive Statistics
	CAATS		Genetic Algorithms	ARIMA	Structural Models
			Expert Systems/ Decision Aids	Univariate and Multivariate	AHP
			Bagging, Boosting		Monte Carlo Study
			BBN		
			Probability Models		
Review:	Ratio Analysis	Visualizations	Expert Systems/ Decision Aids	Linear Regression	Multi-criteria Decision Aid
	CAATS		BBN	Time Series	Descriptive Statistics
			Probability Models	ARIMA	Structural Models
				Univariate and Multivariate	Hypothesis Evaluation
Opinion:	Ratio Analysis	Visualizations	Expert Systems/ Decision Aids	Log Regression	Multi-criteria Decision Aid
				Linear Regression	Descriptive Statistics
Continuous Activities:					

Table 5: Summary listing/draft framework of the techniques occurred in the various Audit Phases in the literature

Based on the analysis of which techniques are used in the various audit phases in the literature, a preliminary mapping (Table 5) is created, based entirely on the discussions in the 301 papers. The predominant techniques for all phases belong to the

Audit Examination and Regression approaches, with some use of BBN, probability models, descriptive statistics, and expert systems. Although it may appear in the framework that many other more complex techniques are analyzed by audit academics, their deployment in the literature is inconsistent and sporadic. Some techniques are discussed only a couple of times, as is the case with text mining, visualizations, process mining, SVM, ANN, Genetic Algorithm, C4.5 Classifiers, AHP, and hypothesis evaluation.

The percentage of papers using specific analytical techniques is shown below in Figure 8. Many papers mention more than one analytical technique. In the realm of audit analytic techniques, the most frequently used techniques are those of Audit Examinations followed by Regressions. Audit Examinations were discussed 459 times; Unsupervised Methods, 43 times; Supervised Methods, 171 times; Regression, 251 times; and Other Statistical Methods, 77 times.

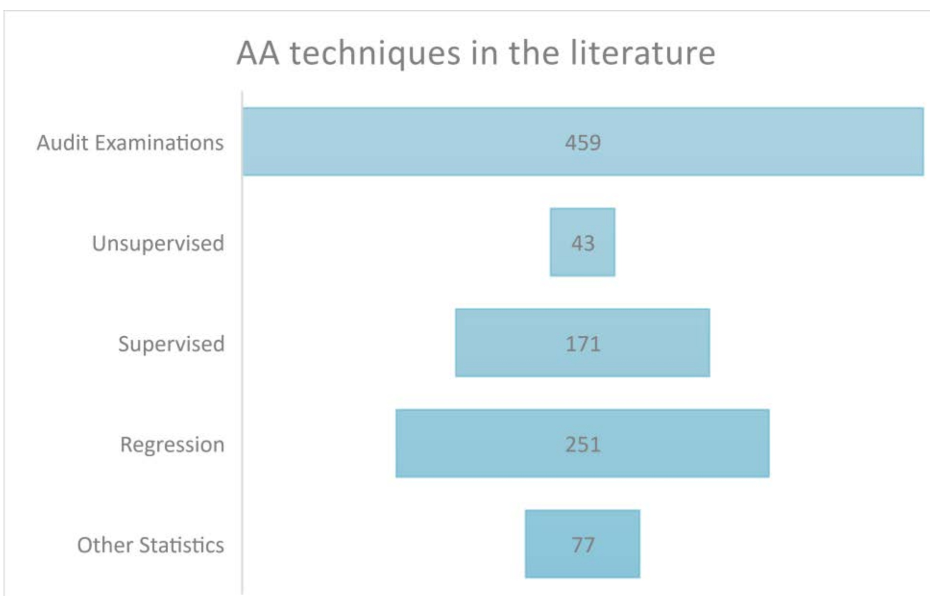

Figure 8: Number of papers using certain Audit Analytics techniques in the literature

In the task of Audit Examination, techniques such as sampling, ratio and trend analysis, CAATS usage, and general ledger tests, there are clear favorites. Sampling techniques and ratio and/or trend analysis are discussed more frequently than any other method, at 37.8% and 43.5% respectively. CAATS are included in this category as many of the tests conducted by external auditors in the papers were general ledger tests and basic calculations (Figure 9).

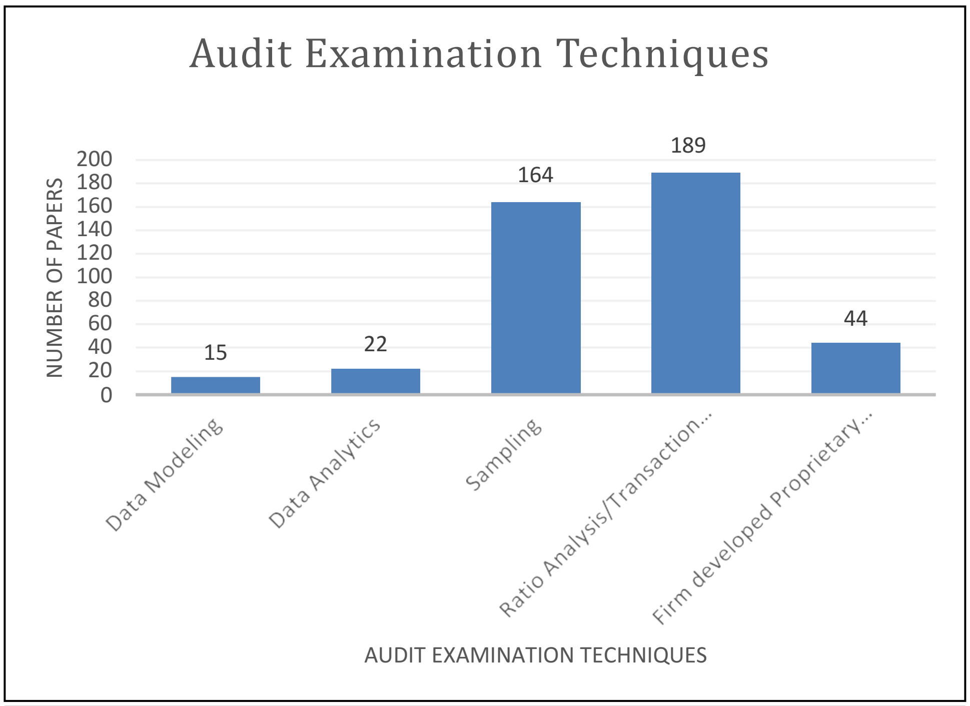

Figure 9: Number of papers that discuss the various Audit Examination techniques

Additionally, Bayesian statistics are applied extensively in the area of sampling (Ijiri & Kaplan, 1971; Corless, 1972; Elliott & Rogers, 1972; Hoogduin, Hall, & Tsay 2010) and in auditor judgment and planning (Felix, 1976; Chang, Bailey, & Whinston, 1993; Dusenbury, Reimers, & Wheeler, 1996; Krishnamoorthy, Mock, & Washington, 1999).

Regression techniques are second in popularity, discussed 251 times in the audit literature. Log Regression was mentioned 81 times, with Linear Regression at 62 times, Time Series Regression at 34 times, ARIMA at 20, and Univariate and Multivariate at 54 (Figure 10).

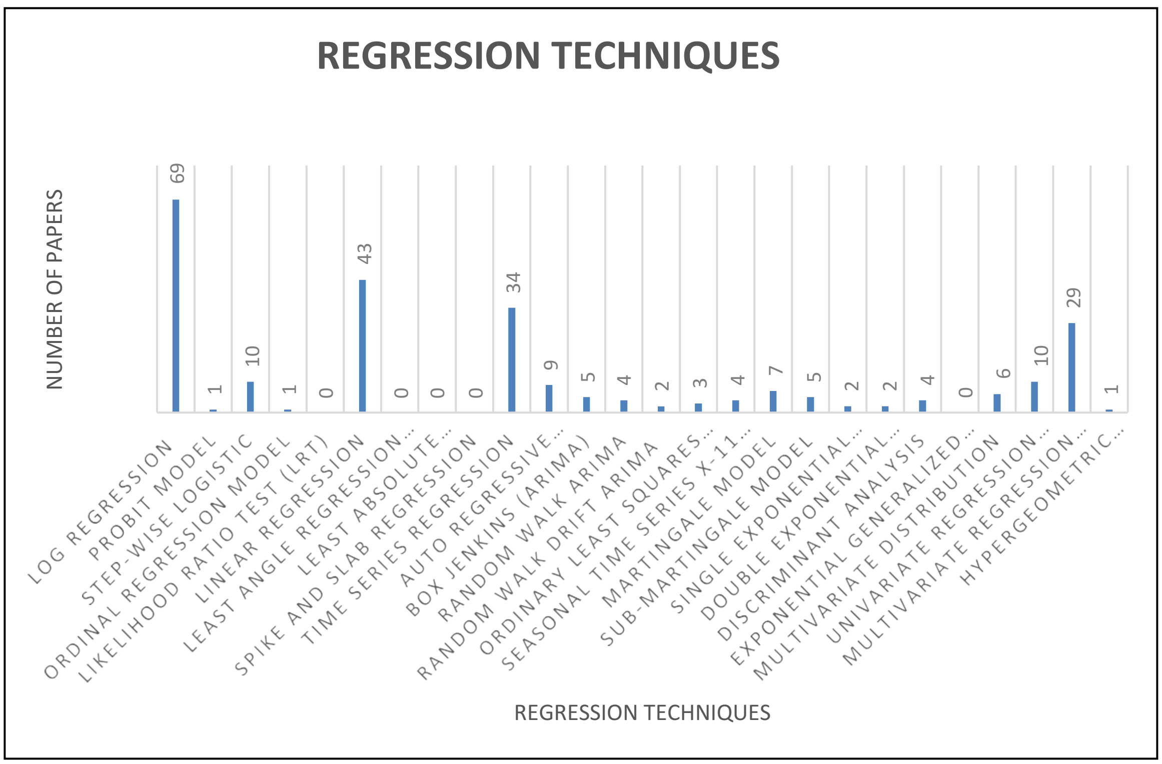

Figure 10: Number of papers discussing Regression Methods

Most popular of the supervised techniques is the application of Bayes Learners/Bayesian Belief Networks at 46 times, followed by Expert Systems at 41, Probability Models at 30, and Artificial Neural Networks at 24 times (Figure 11).

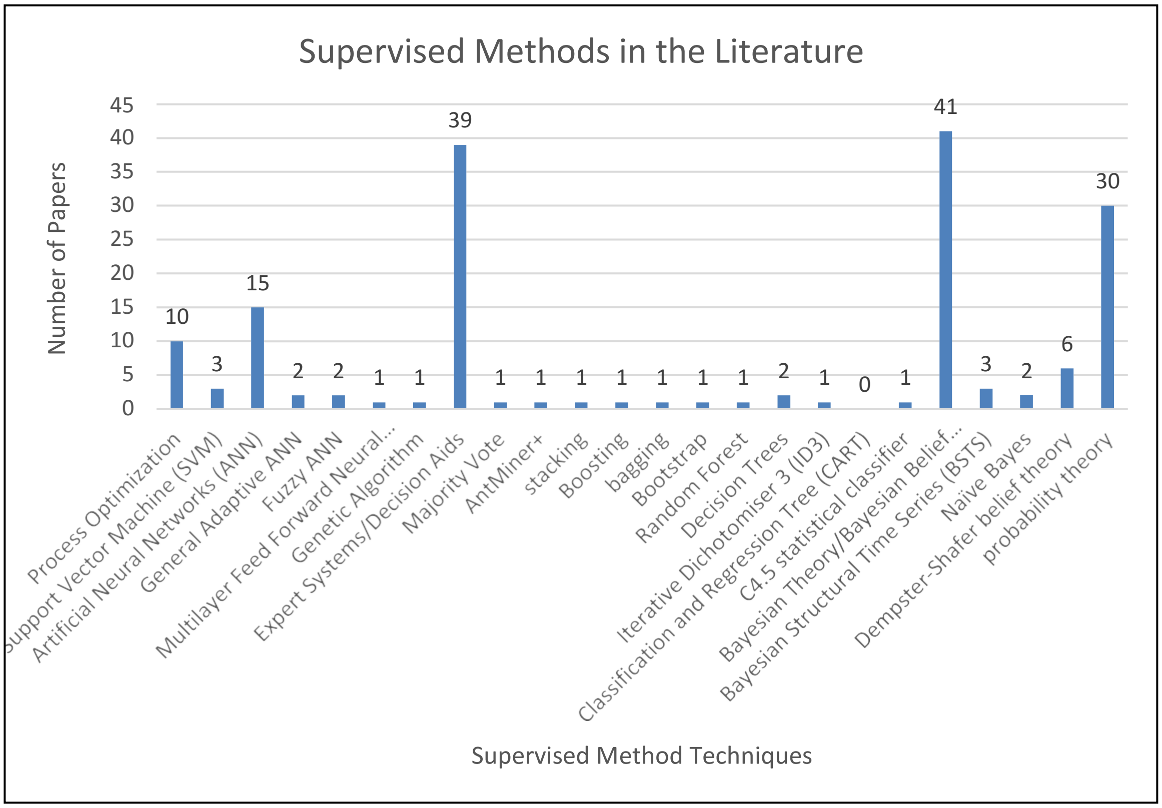

Figure 11: Breakdown of Supervised Methods by technique and the number of times each is discussed

Unsupervised Methods are discussed minimally, with Process Mining being the most popular (Figure 12).

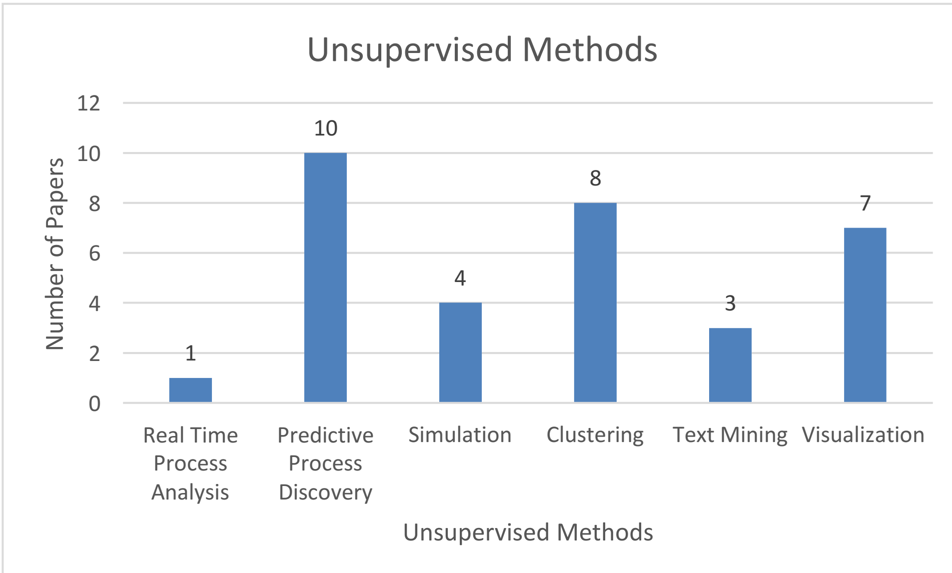

Figure 12: The number of papers discussing each Unsupervised Method

Other Statistical Methods are slightly more popular with coverage in 77 papers, with Descriptive Statistics receiving the most attention in 31 papers (Figure 13).

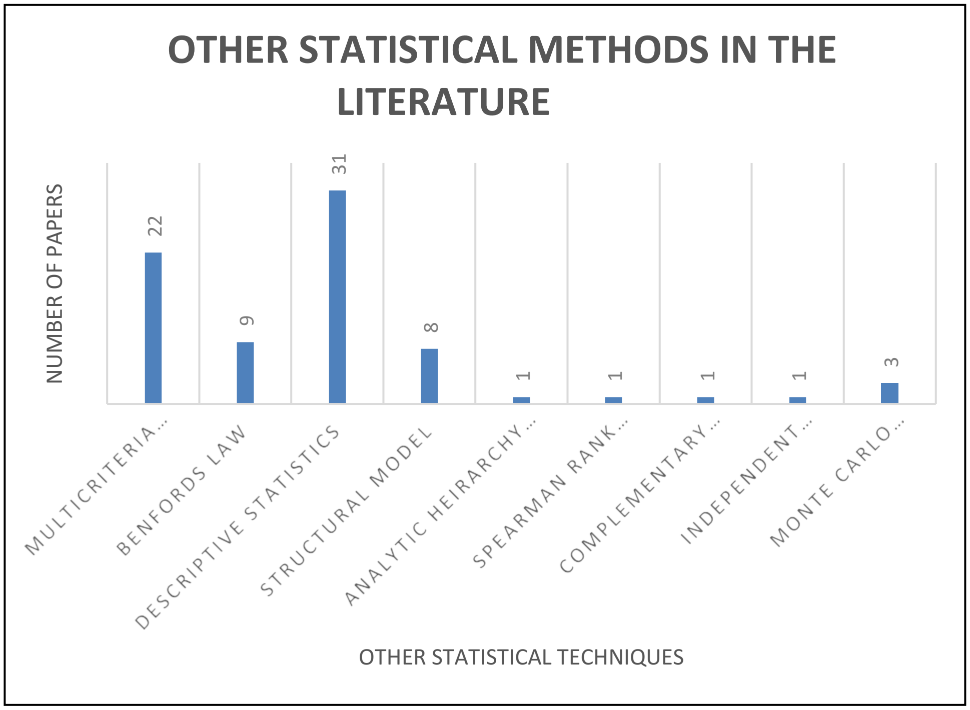

Figure 13: The number of times that Other Statistical Methods are discussed

2.3 Evolution of the External Audit Analytics Framework

2.3.1 (RQ3) How to organize the main attributes covered by these studies of analytical procedures in the audit engagement?

The sheer number of papers still presents a challenge for researchers even after many features have been described. The systematic research method (Keele 2007) suggests that an organizing conceptual framework should be developed to facilitate understanding. The aim of this structured research is not just to aggregate the evidence but to also provide guidelines for future academic research and practitioner applications in a specific context.

A conceptual framework may be defined as “the way ideas are organized to achieve a research project’s purpose” (Shields and Rangarjan 2013, p 24). For RQ3, the purpose of a framework is to organize the literature to best understand how researchers apply analytical procedures to the audit engagement. Since the typical engagement proceeds with the format of the audit phases, it seems logical to organize the literature first by audit phase and then these phases are subsequently divided by AP type. Table 5 summarizes this information which is presented in detail with paper numbers in Table 21 of Appendix A. The numbers are assigned for each paper in Table 9 of Appendix B. However, Table 21 with its lists totaling 301 papers may still appear overwhelming. Therefore, it may be appropriate to organize this literature within another view of APs, that of Business Analytics (BA).

2.3.1.1 Business Analytics

Since auditors examine business financial data, much of which may be generated with applications and analytics embedded in management enterprise systems, gaining knowledge of and perhaps adapting concepts of business analytics (Holsapple et al, 2014) could be beneficial. Business analytics is 'the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their operations, and make better, fact-based decisions' (Davenport and Harris, 2007). The recently proposed three dimensions of domain, orientation, and techniques (Holsapple et al 2014) are useful for understanding the scope of business analytics. Domain refers to the context or environment in which the analytics are being applied. Orientation describes the outlook of the analytics – descriptive, predictive, or prescriptive, while techniques refer to the analytical processes of the domain and orientation. The feasibility of the application of a certain technique is not only dictated by its orientation, but also by the available data.

In the environment that the audit team operates, the domain dimension of the client is business enterprise and management. The three dimensions of orientation should be clarified to gain an understanding of their roles in the business domain. The differing orientations of these dimensions are partly due to the availability of different types of data in conjunction with various techniques and the capabilities of the client enterprise systems.

Descriptive Analytics

Descriptive analytics answers the question as to what happened. It is the most common type of analytics used by businesses (IBM, 2013) and is typically characterized by descriptive statistics, Key Performance Indicators (KPIs), dashboards, or other types of visualizations (Dilla, Janvrin, and Raschke 2010). Descriptive analytics also forms the basis of many continuous monitoring alert systems, where transactions are compared to data based analytics (Vasarhelyi and Halper 1991) and thresholds are established from ratio and trend analysis of historical data.

Predictive Analytics

Predictive Analytics is the next step taken with the knowledge acquisition from descriptive analytics (Bertsimas and Kallus, 2014) and answers the question of what could happen (IBM, 2013). It is characterized by predictive and probability models, forecasts, statistical analysis and scoring models. Predictive models use historical data accumulated over time to make calculations of probable future events. Most businesses use predominantly descriptive analytics and are just beginning to use predictive analytics (IBM, 2013).

Prescriptive Analytics

Prescriptive Analytics (Bertsimas and Kallus, 2014; Holsapple et al, 2014; IBM, 2013; Ayata, 2012) answers the question of what should be done given the descriptive and predictive analytics results. Prescriptive analytics may be described as the optimization approach. Prescriptive analytics go beyond descriptive and predictive by recommending one or more solutions and showing the likely outcome of each.

The techniques for predictive and prescriptive analytics may appear similar, but their orientation and ability to prescribe or predict depends on the type and amount of data available for analysis. The bigger the data and more varied the data types, the more likely the solution may be prescriptive. Prescriptive techniques may pull upon quantitative and qualitative data from internal and external sources. Analytics based on quantitative financial data alone are utilizing only a fraction of all available data, since most data is qualitative (Basu, 2014). Based on business rules, constraints, and thresholds, in a prescriptive orientation, mathematical simulation models or operational optimization models are built that identify uncertainties and offer solutions to mitigate the accompanying risks or adverse forecasts (Appelbaum et al 2016).

The techniques of business analytics can be considered as either qualitative or quantitative, or as deterministic or statistical, or based on unstructured, semi-structured, or structured data (Table *20* in Appendix A). The most traditionally used accounting techniques are those that are quantitative, statistical, and based on structured data. While in the past most advanced business analytics techniques came from statistical data analysis, more recently research has begun incorporating techniques that originate in machine learning, artificial intelligence (AI), deep learning, text mining, and data mining. Some of these techniques do not make any statistical assumptions about underlying data, and consequently generate models that are not statistical in nature. The techniques found in business analytics are classified in Table *20* located in Appendix A.

Because the standards refer to analytics as "analytical procedures" (APs), this research refers to the use of any type of analytics in the audit literature as APs. When discussing these techniques in a context outside of the literature, the terminology will be that of analytics or business analytics. Given the attributes of APs as discussed in the literature, the next challenge is to obtain an understanding of how APs can relate to Business Analytics. This process starts by first understanding the literature to date, by undertaking the next steps of the literature review process.

2.3.2 (RQ4) Given the attributes categorized in RQ3, how can this literature be presented to direct future research?

One of the more common reasons for performing Systematic Literature Review (SLR) is to provide a framework or context to appropriately position new research activities, having identified the extant research (Keele 2007, p 3). Within this scope of SLR exists the possibility of a Systematic Mapping Study (SMS) (Keele 2007 p 44). SMS provides a broad overview of the literature with the intent to influence the direction of future research. The analysis stage of an SMS oriented SLR summarizes the data to answer the research questions. These data summaries are then disseminated by means of SMS conceptual framework. Ideally, this SMS conceptual framework should be impactful to current practice and developments.

This paper began by describing the dilemma of the current audit profession, that the emergence of big data as well as the growing use of analytics by audit clients has brought new concerns. That is, audit clients are progressively using more complex Business Analytics (BA) and auditors are concerned that APs as typically and historically applied may not be effective. Since auditors examine business financial and BA data, ideally a SLR/SMS based framework should reflect these new concerns. To maintain relevancy, current audit academics should examine those areas that are lacking research to date.

This section will discuss the evolution of a conceptual External Audit Analytics (EAA) framework, based on this examination of extant audit academic research within the more general context of Business Analytics (BA). Although there have been many applications of basic analytics in the external audit practice[3] there should be a framework providing guidance for research of the more complex analytical techniques. With this proposed framework, it is hoped that academics will feel more comfortable expanding the scope and nature of their research about analytics in the audit. External Audit Analytics (EAA) or analytical procedures comprises the utilization of various analytical methods and models to facilitate the transformation of data into external audit evidence and subsequently into audit decisions.

[3] Li et al.(2016) surveyed users of an audit analytics software and found very limited use of advanced analytics.

EAA may be considered as a special sub-area of the wider area of Business Analytics (BA) since public auditors examine business financial data. Business Analytics is discussed in the previous section and its features are subsequently applied to the Analytical Procedures function of the audit engagement. In this context, APs as practiced to date (Table 1) are but one component of EAA. APs in the context of EAA are much more than the APs as conventionally understood (Table 1). The conventional Analytical Procedures (APs) process, when regarded under the view of Business Analytics, can now be conceptually regarded as a component of External Audit Analytics (EAA). For example, in Table 1 APs are described as basic comparisons and ratio analysis using both financial and nonfinancial data – however, EAA pertains to all BA techniques that lend themselves to the engagement process.

Accordingly, the discussed above three BA dimensions (Holsapple et al, 2014) are useful for conceptualizing EAA as well. EAA and BA may appear to be quite similar. The advances observed in the extant literature to date will be categorized by these dimensions of domain, orientation, and technique. These dimensions, particularly that of orientation, are a new way of understanding analytics in the external audit. The mapping of BA to EAA to conventional APs is illustrated in

Figure 14. Here APs are depicted as a subset of EAA, which is a subset of BA. Figure 14 represents the current literature based research, where Figure 15 imagines where the literature could evolve in the future. Figure 15 illustrates how the circle of understanding that is conventional APs could be expanded to include all conceptual EAA techniques. Conceivably, APs could now equal EAAs, and as such could possess the dimensions of EAA, such as orientation of descriptive, predictive, and prescriptive.

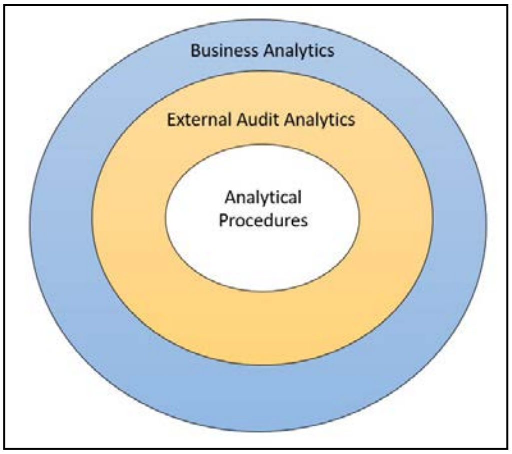

Figure 14: Mapping of conventional Analytical Procedures, where they are shown to be a subset of EAA which is a subset of BA

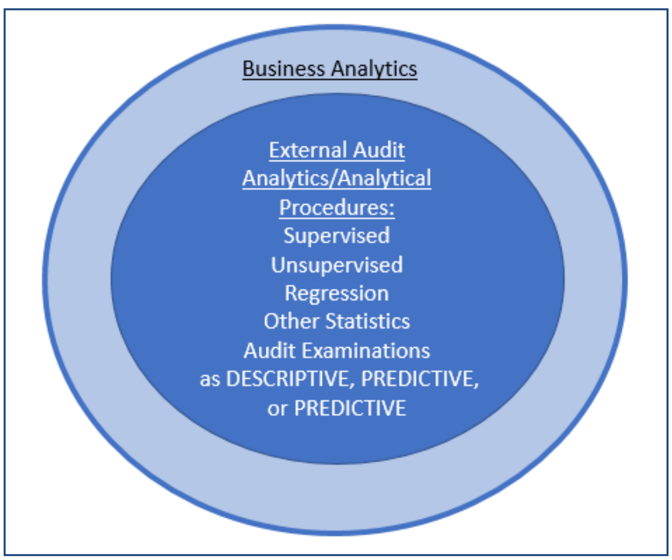

Figure 15: Mapping of the Conceptual EAA which now may be understood interchangeably as APs.

Using the dimension of orientation (descriptive, predictive, prescriptive) to assist in forming a literature based framework, a process flow for understanding and categorizing the 301 papers can now be established (Figure 16):

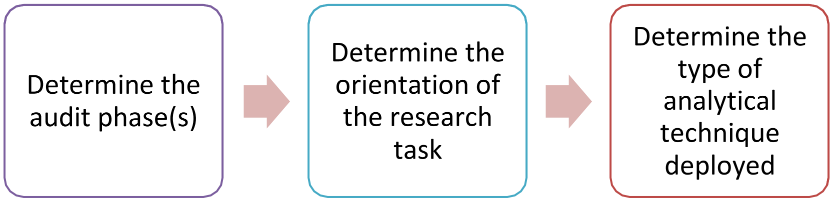

Figure 16: Process Flow of the Literature Framework Formation

What follows (Figure 17) is a high level graphic illustration of the literature-based framework which identifies all of the APs that the papers discuss:

Engagement

Orientation:	Audit Examinations	Unsupervised	Supervised	Regression	Other Statistics
Descriptive	√	√			√
Predictive			√	√	√
Prescriptive					

Planning/Risk Assessment

Orientation:	Audit Examinations	Unsupervised	Supervised	Regression	Other Statistics
Descriptive	√	√			√
Predictive			√	√	√
Prescriptive					

Substantive & Compliance Testing

Orientation:	Audit Examinations	Unsupervised	Supervised	Regression	Other Statistics
Descriptive	√	√			√
Predictive			√	√	√
Prescriptive					

Review

Orientation:	Audit Examinations	Unsupervised	Supervised	Regression	Other Statistics
Descriptive	√	√			√
Predictive			√	√	√
Prescriptive					

Opinion Formulation and Reporting

Orientation:	Audit Examinations	Unsupervised	Supervised	Regression	Other Statistics
Descriptive	√	√			√
Predictive			√	√	√
Prescriptive					

Continuous Activities

Orientation:	Audit Examinations	Unsupervised	Supervised	Regression	Other Statistics
Descriptive					
Predictive					
Prescriptive					

Figure 17: The scope of Analytical Procedures as covered in the literature in the conceptual domain of EAA

Figure 17 categorizes the literature at a high summary level and could be regarded as the literature based framework for APs in the external audit domain.[4] Checkmarks indicate where a paper has been identified, based on the process in Figure 16 for that phase/orientation/technique type. This table portrays the attributes of all the literature to date with the "new" understanding of orientation. All the blank spaces represent areas where literature has not been found to date yet are potential areas of research in BA. Conceptually, the blank spaces represent areas of unexplored EAA in extant research. However, before assuming all blank spots automatically represent EAA, a better understanding of EAA would be helpful.

2.4.2.1 Domain of EAA

The domain of external auditing is naturally associated with the stages of the audit cycle where EAA methods and models may be applied. Issues which may emerge during this process could be as follows:

- How different are the objectives of Internal and External Audit Analytics in the current context (Li et al, 2016)?

[4] It is recommended that interested researchers follow these procedures:

- ✓ First, identify the area(s) of interest in Figure 17 here
- ✓ Secondly, look at those phases and their more detailed AP type in Table 20 to obtain more insight about the techniques
- ✓ Thirdly, look at Table 21 in Appendix A under the specific AP technique(s) and phase(s) (from Table 20) to gather all relevant paper numbers
- ✓ Finally, find these paper numbers in Table 9 (Appendix B) for research and analysis

- Isn't there a substantive overlap between business monitoring and real time assurance?
- Considering that there is substantive overlap in data analytic needs, are the traditional three lines of defense (Freeman, 2015; Chambers, 2014)) still relevant?

2.4.2.2 Orientation of EAA

A distinction can be drawn regarding descriptive, predictive, and prescriptive orientations of EAA. These are discussed earlier but are quickly reviewed here:

Descriptive EAA answers the question as to what happened. It is the most common type of analytics used by auditors and is typically characterized by descriptive statistics, Key Performance Indicators (KPIs), dashboards, or other types of visualizations.

Predictive EAA is the next step taken with the knowledge acquisition from descriptive analytics (Bertsimas and Kallus, 2014) and answers the question of what could happen (IBM, 2013) and is characterized by predictive and probability models, forecasts, statistical analysis and scoring models. Most audit clients use predominantly descriptive analytics and are just beginning to use predictive analytics (IBM, 2013). The following issues should be considered by audit researchers in this evolving analytic environment:

- Traditional auditing has a retrospective approach, as traditional technologies did not allow for other approaches - can the current environment allow for a prospective look?

- What parts / procedures of the audit are fully or partially automatable?
- Will it allow a disruptive change (Christensen, 2013)?

Prescriptive EAA (Bertsimas and Kallus, 2014; Holsapple et al, 2014; IBM, 2013; Ayata, 2012) goes beyond descriptive and predictive by recommending one or more solutions and showing the likely outcome of each It is a type of predictive EAA in that it prescribes a solution requiring a predictive model with two components: actionable big and varied (hybrid) data and a validation/feedback system. A prescriptive EAA model will have a decision function that chooses among alternatives. Interesting questions emerge from attempting to prescribe:

- Can the key contingencies in the audit be formalized?
- Will these be allowed to evolve under the current audit standards?
- Are they so disruptive (Christensen, 2013) that they will be ignored by current leading audit firms?

2.4.2.3 Techniques of EAA

EAA undertaken in an engagement where big data is available may result in a prescriptive analytics approach where a set of techniques computationally identifies several alternative actions to be taken by the auditor, given the audit's complex objectives and limitations, with the goal of reducing audit risk. For example, EAA techniques utilizing varied sources of big data could be used to arrive at a quantitative score for the audit opinion, as opposed to the current pass/fail opinion.

The currently mandated pass/fail opinion formatdoes not reflect the nuances and details of the auditor's work - the culmination of much laborious examination and careful judgement by the auditor. With more advanced EAA techniques and reliable evidence, it is probable that this process and resulting opinion could be quantified with prescriptive analytics. Prescriptive analytics may allow for a graduated scale or ranking of audit opinion and audit risk. In an ideal scenario, auditors should be prolific in their use of analytic techniques of all three orientations, as analytics should be dominant in industries that are very data-rich and where one of the major improvements from analytics usage is risk reduction (Banerjee et al, 2013).

2.4.2.4 The Integration of the Literature Framework with EAA

Many of the techniques observed in the external audit literature are quantitative in nature. This dominance of quantitative techniques in APs may be because the main objective of external audit has been to provide assurance on the accounting numbers. Therefore, the accounting numbers traditionally were the focus of APs. However, with the availability of social media and big data, the scope of APs could be expanded to that of EAA. This greater variety of available data creates the opportunity for more advanced analytics research.

Accounting numbers are derived by manipulating (aggregating, adjusting, etc.) quantitative descriptions of business transactions that are currently typically stored in relational tables. Such data are obviously well structured. These structured data lead typically to analysis which is quantitative and descriptive, and can be categorized as Audit Examination techniques. Audit Examination entails, among many procedures, basic transaction tests, three-way matching, ratio analysis, sampling, re-confirmation, and re-performance. These tests are applied in every external audit engagement, and are regarded as fundamental EAA. These tests may be performed manually or with the assistance of Computer Assisted Auditing Tools (CAATs). Sampling in this context may be statistical or non-statistical and of attribute or monetary type – however, it is categorized as fundamental audit examination in EAA due to the pervasiveness of the technique (Elder et al, 2013).

2.4.2.5 Expectation Models and EAA

To obtain the context for how the EAA framework could fit in an audit engagement, a formal discussion of expectation models in the audit is required. The most common types of techniques utilized in EAA, in addition to those of audit examination, are expectation models. A typical expectation model is an empirical relationship among several accounting numbers or some other important quantitative measures of business operations. Such relationships hold only in the statistical sense, up to certain error terms, that are usually assumed to be random.

An expectation model is inferred from the archive of historical records. If it turns out to be possible to infer a stable empirical relationship that fits the historical records well, then it is reasonable to expect this relationship to hold in the near future, assuming no significant changes take place in the business. Therefore, this relationship provides an expectation model for the accounting numbers and other important business metrics of the near future. The accuracy of this future relationship provides important audit evidence about the veracity of the quantities involved.

It is common to focus on a certain accounting number (e.g., revenue), and represent an expectation model as an equation for this accounting number. Then, for a given confidence level, this equation can be used to derive a prediction interval for the future value of the accounting number. If the actual future value turns out to be inside the prediction interval, this can be interpreted as strong evidence that the accounting number is properly represented. Otherwise, the auditor will need to conduct further investigation to determine if there is indeed a problem with this accounting number. The expectation model forms the basis of audit examination in the engagement and determines the direction and degree of evidence collection and audit scrutiny.

The EAA usage described above has predictive orientation, and the amount of audit evidence provided is based on the level of agreement between the observed business reality and the predictions. This is utilized not only to verify accounting numbers, but also to provide assurance on controls by comparing the observed business process workflow with the expectations derived either from the existing business rules, or from the past observations of business processes. As an example of the former, a business rule stating that "purchase orders exceeding $1,000 require management authorization" creates an expectation with which all future purchase order transactions would be compared. As for the latter option, if the analysis of past purchase orders shows that 99% used vendors that were pre-approved, then it would be reasonable for the auditors to expect that every future purchase order would use a pre-approved vendor, and those that do not would warrant investigation.

2.4.2.6 EAA Expanded

While the EAA expectation models that have been derived by formalizing business rules are usually essential in the current engagement domain, they are not as methodologically challenging and this manuscript focuses on other EAA expectation models obtained from more advanced techniques.

The most basic dichotomy of the EAA techniques distinguishes between structural and quantitative methods. Structural techniques look for various structural properties in the historical records. A recent example is process mining (Jans et al, 2013). It provides techniques for analyzing enterprise system logs and identifying the most common paths of enterprise business workflow to be used as expectation models. If the observed workflow of a particular process deviates significantly from the expected path, it should warrant an investigation.

In the realm of quantitative techniques, it is appropriate to make a distinction between univariate and multivariate methods. Univariate techniques usually infer various distribution properties of individual quantities, and can be as familiar as estimating the median, mean, skewness and kurtosis, or more complex as applying Benford’s law to auditing.

There is a great variety of EAA multivariate techniques, and no generally accepted agreement on their taxonomy[5]. It could be useful to differentiate multivariate

[5] The primary objective of multivariate techniques is to develop relationships between or among variables/features under study. In this view, the universe of multivariate techniques is wider than what is usually considered to be the domain of multivariate statistics, where joint distributional properties of more than one variable are studied. If only a single variable is viewed as the outcome or dependent variable, and its univariate distribution is studied given the values of some the other variables, such as case in

techniques by considering whether a particular EAA technique explicitly assumes the presence of latent[6] features. For example, common classification and regression techniques do not work explicitly with any latent features, while common clustering techniques do (with the latent feature being the cluster ID). Often, the utilization of latent features techniques is necessitated by the lack of critical information in the historical records. For example, while it is commonly assumed that managerial or financial statement fraud is a routine occurrence in most enterprises, very few confirmed and documented cases of such fraudulent transactions exist. For this reason, most audit engagement teams face the challenge of creating expectation models for what is fraudulent versus normal, given that the historical records do not identify past transactions in this way.

Another important technique dimension to consider is the scale of variables utilized in the expectation models, with the categorical and continuous ones being the two most commonly used general types. The two important measurement scales of categorical variables are nominal and ordinal, while the two important measurement scales of continuous variables are interval and ratio.

It is often the case that a technique assumes that all the variables are measured on one type of the scale, and adaptations are required for those measured on a different one. For example, multiple linear regression models are developed for the case of

---

multiple linear regression, then we view it as a multivariate technique even though it is traditionally not considered to be multivariate statistics.

[6] Latent features are attributes or qualities that are not directly observed. For example, a concept such as trust is measured in terms of multiple indirect observations that have shown correlation with it, thereby deriving a number for this attribute which cannot be directly measured.

continuous variables, while the categorical scales of independent variables are accommodated by using dummy variables. Sophisticated generalizations of multiple linear regression models such as ordinal regression models are utilized to deal with the case of categorical dependent variables. On the other hand, decision trees are developed for nominal variables, while the continuous ones are accommodated by introducing their comparisons with threshold values.

An important subset of continuous EAA models consists of the time series models, where the time variable is afforded special treatment. Note that univariate time series models are based on two variables (including time). Also, commonly used time series models study relationships between variable values at discrete moments in time. Those much more complicated models where time is continuous belong to the realm of stochastic processes, and such models have not so far found applications in audit analytics.

2.4.2.7 The EAA Framework

Combining knowledge of the EAA with the literature framework of Figure 17, a summary conceptual framework of audit analytics for the external audit domain is proposed (Figure 18). This EAA framework satisfies the objective of RQ4, in that it provides a guideline for future research in the domain. By grounding the EAA framework with analytics based on prevalent business and external audit practices, future research maintains its relevance to the profession.

This framework identifies those areas of APs (now considered as EAA) that have been covered by extant literature and those areas of research that exhibit gaps in the

EAA domain. By utilizing the literature supported framework from Figure 17, each of the audit phases of Engagement, Planning/Risk assessment, Substantive & Compliance Testing, Review, Opinion Formulation and Reporting, and Continuous Activities could be enhanced with EAA in as follows in Figure 18:

Engagement: The auditors have access to the audited financial statements and other public information as well as other external sources of data, not dissimilar to investment/financial analysts. It is envisioned that auditors could assess the desirability of engaging/retaining a client using many of the analytic techniques that are undertaken by most financial analysts. Expectation models could be developed at this time, derived from quantitative and qualitative data. At this stage, auditors could perform the following techniques: ratio analysis of audited statements, text mining, visualization, expert systems, belief networks, probability models, regression, and descriptive statistics.

1. Planning/Risk Assessment: Similar to the Engagement Phase, but the auditors now have access to the current unaudited financial statements and can develop models of what could and should happen. Clustering, visualization, regression, belief networks, expert systems, and descriptive statistics may be used in addition to ratio and trend analysis.

2. Substantive Testing & Compliance Testing: This phase could entail sampling as well as testing of 100% of the transactions, depending on the client environment. Transactions could be tested against benchmarks and expectation models. Results

that are flags or indicative of further investigation could be subject to further testing and evidence collection. However, initially this phase most likely would include all audit examination techniques, Audit by Exception (ABE) if appropriate, clustering, text mining, process mining, visualization, SVM, ANN, expert systems, decision trees, probability models, belief networks, regression, Benford's Law, descriptive statistics, structural models, and hypothesis evaluation.

3. Review: This phase could entail cross-validation tests and analysis of exceptional results using different techniques. This phase will lean more towards prescriptive testing, as what should have happened will serve as the benchmark of what happened. All the techniques outlined in Substantive Testing could be applied here, with more emphasis on expert systems, probability models, belief networks, SVM, ANN, genetic algorithms, multi-criteria decision aids, regression, and hypothesis testing.

4. Opinion Formulation and Reporting: This phase is based on the comparison between what could and should happen and what actually happened, and the greater the difference between these two expectations, if not corrected by the client, the more likely a qualified opinion. It is anticipated that there may be a more nuanced measurement of risk than the current unqualified/qualified opinion. Potentially the audit opinion could be a more informative, graduated opinion derived from prescriptive analytics of reliable evidence. This phase could feasibly benefit from the same approaches mentioned in earlier phases, with more emphasis

on time series regression, probability models, belief networks, expert systems, and Monte Carlo studies. The topic of the application of analytical techniques to arrive at a more quantitative audit opinion, away from the current mainly dichotomous outcome, is an area for future research.

5. Continuous Activities: The auditor may run continuous or interim tests using many different models to generate predictive and prescriptive expectations of the ongoing client's activities and how they may impact the upcoming financial statements. This phase would involve the use of many audit examination techniques as a foundation for the use of regression, descriptive statistics, belief networks, probability models, expert systems, decision trees, process mining, visualization, text mining, and clustering. Prescriptive models would be continuously updated with new data, improving the models' accuracy over time. Continuous Auditing (CA) (Vasarhelyi and Halper 1991) with its real-time feed of relevant information could be considered as an interim continuous activity.

Audit Examination techniques form the foundation of each step in the proposed EAA framework. Since Audit Examination techniques may be descriptive, exploratory, and confirmatory (Liu 2014), they provide a level of domain and transaction knowledge that are essential to the auditor. In EAA, it is expected that data preparation procedures such as data verification, data cleaning, and data harmonizing contribute to "client knowledge" or "client data expertise" and are similarly time-consuming and laborious to obtain.

The framework in Figure 18 displays the general type of technique (Audit Examinations, Unsupervised, Supervised, Regression, and Other Statistics) that potentially could be deployed by auditors and the orientation of these techniques (Descriptive, Predictive, and Prescriptive). This framework may serve as a foundation for additional detailed research by practitioners, standard setters, and academia regarding the use of the various suggested techniques for each audit phase. The areas that are checked without stars are areas that are being researched already (Figure 18). The phases where research appears to be missing to date or is scant are highlighted with stars. The gaps shown in Figure 18 are identified now as research-sparse EAA. For example, clustering as an unsupervised descriptive method has been found to be missing in the engagement phase literature and is suggested here for future analysis. Or, visualization as an unsupervised method has been examined for many audit phases in some research; however, this does not mean that there isn't room for additional research contributions. In general, the phases of Engagement, Opinion, and Continuous Activities are particularly sparse, most likely since the standards do not require analytical procedures at these phases and therefore could benefit from additional research.

The proposed EAA framework is based on the assumptions that the auditor has few technical constraints and has access to a significant amount of client and other external data. Figure 18 combines the discussion of the potential approaches for possible technique types in each audit phase (see beginning of this section) with that of the literature framework (Figure 17).

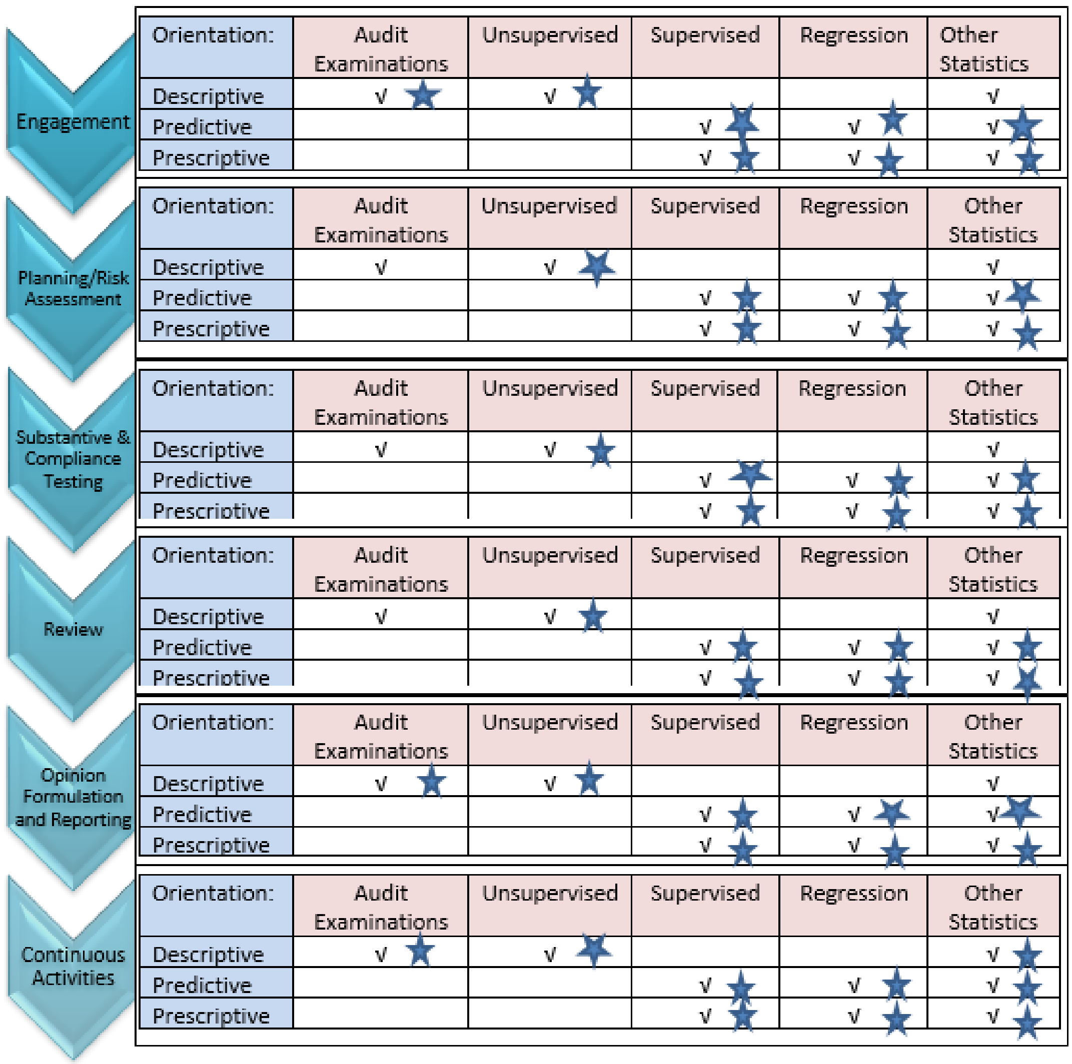

Figure 18: Conceptual External Audit Analytics (EAA) Framework

Figure 18 proposes that at least one of each technique types from Audit Examination, Unsupervised, Supervised, Regression, and Other Statistics could be undertaken in each phase of the external audit. There are research gaps in visualization, process mining, and all prescriptive methods for every audit phase. In this model, due to the increased availability of many types of internal and external data, analytics may be used in every phase of the audit. The implementation of any EAA technique for a certain phase would depend on the audit objective and relevant assertions. Figure 18 motivates Table 6 which displays the potential for each technique and orientation per audit phase to be explored in future research:

Descriptive	Engagement	Planning	Testing	Review	Opinion	Continuous activities
Clustering Models	✓	✓	✓	✓	✓	✓
Descriptive Statistics						✓
Process Mining: Process Discovery Models	✓	✓	✓	✓	✓	✓
Ratio Analysis						✓
Spearman Rank Correlation Measurement		✓	✓	✓		✓
Text Mining Models			✓	✓	✓	✓
Visualization	✓	✓	✓	✓	✓	✓
Predictive	Engagement	Planning	Testing	Review	Opinion	Continuous activities
Analytical Hierarchy Processes (AHP)	✓	✓		✓	✓	✓
Artificial Neural Networks (ANN)	✓	✓		✓	✓	✓
Auto Regressive Integrated moving Average (ARIMA)					✓	✓
Bagging and Boosting models	✓	✓		✓	✓	✓
Bayesian Theory/Bayesian Belief Networks (BBN)	✓				✓	✓
Benford's Law	✓	✓		✓	✓	✓
C4.5 Statistical Classifiers		✓	✓	✓	✓	✓
Dempster-Shafer Theory Models	✓	✓	✓	✓	✓	✓
Expert Systems/Decision Aids	✓					✓
Genetic Algorithms	✓	✓		✓	✓	✓
Hypothesis Evaluations	✓	✓	✓		✓	✓
Linear Regression	✓	✓				✓
Log Regression		✓		✓		✓
Monte Carlo Study/Simulation	✓	✓	✓	✓	✓	✓
Multi-criteria Decision Aid				✓		✓
Probability Theory Models	✓				✓	✓
Process Mining: Process Optimizations	✓	✓	✓	✓	✓	✓
Structural Models					✓	✓
Support Vector Machines (SVM)	✓	✓		✓	✓	✓

Time Series Regression					✓	✓
Univariate and Multivariate Regression Analysis					✓	✓
Prescriptive	Engagement	Planning	Testing	Review	Opinion	Continuous activities
Artificial Neural Networks (ANN)	✓	✓	✓	✓	✓	✓
Auto Regressive Integrated Moving Average (ARIMA)	✓	✓	✓	✓	✓	✓
Expert Systems/Decision Aids	✓	✓	✓	✓	✓	✓
Genetic Algorithms	✓	✓	✓	✓	✓	✓
Linear Regression	✓	✓	✓	✓	✓	✓
Log Regression		✓	✓	✓	✓	✓
Monte Carlo Study/Simulation	✓	✓	✓	✓	✓	✓
Time Series Regression	✓	✓	✓	✓	✓	✓
Univariate and Multivariate Regression Analysis	✓	✓	✓	✓	✓	✓

Table 6: Gaps and Areas of Scant research in the EAA context (adapted from Appelbaum et al 2016)

For example, an unsupervised technique such as Visualization which is already predominant in BA (Holsapple et al, 2014) might be readily accepted to supplement audit examination techniques in each phase. It is anticipated that techniques that are of descriptive orientation (audit examination, unsupervised, and other statistics) would be employed first for EAA as these are like audit examination in that they are descriptive. Techniques that are of predictive orientation (unsupervised, supervised, regression, and other) would be next, followed by prescriptive oriented techniques (unsupervised, supervised, regression and other).

As it stands now, auditors typically face significant challenges to obtain sufficient and reliable client evidence. Looking forward, it is believed that these assumptions regarding the EAA framework are not unrealistic – many clients today process dozens of terabytes of internal data, not to mention acquiring additional external sources of data, which is more than a 1000 times the data available just ten years previously (Banerjee et al, 2013). Over time, clients may expect deeper insights from their external auditors, to maximize the potential benefits of their investment in internal IT infrastructure and big data collection. Other client stakeholders may also expect deeper levels of analysis from the external auditor in this big data technology driven business environment.

By and large most advanced analytical procedures are of value for predictive methods but not necessarily prescriptive. Descriptive methods complement these approaches. Traditional descriptive methods can also be supplemented by other statistical methods. This huge potential usage of predictive and prescriptive methods also raises the issue of the adequacy of the traditional organization of the audit in an assurance process that is close to real time, mainly automated, subject to deep human decision making, and complemented by analytic technology.

2.6 Concluding Comments

This research is motivated by the first research issue stemming from the current demands of academia, regulators, and the profession for guidance regarding the increased use of analytics in external auditing. Upon exploration of the academic audit literature for such guidance, it appears that a comprehensive and updated synthesis does not exist. Accordingly, the vast body of audit literature is searched for those papers that discuss the use of analytics in at least one phase of the external engagement. This literature is then examined and categorized by audit phase, analytic technique, orientation, and other details. This literature is then organized into a draft framework which lists the techniques discussed in each phase as proposed by Cushing-Loebbecke (1986).

The literature-based framework is then expanded with the concepts of business analytics (Holsapple et al. 2014), applications which capture the potential information made possible with big data. The revised draft literature framework, now called the External Audit Analytics (EAA) framework, is organized around descriptive, predictive, and prescriptive orientations. Although predominantly literature based, the EAA framework contains recommendations for the utilization of prescriptive techniques.

This chapter organizes and synthesizes the previously uncategorized extant literature, thereby encouraging further research and exploration by academia, regulators, and practitioners. However, due to the very large number of publications, the process of organizing and understanding this research is just beginning. Papers that discuss techniques may report negative findings. For example, even though ratio analysis was and continues to be a predominant EAA technique for practice, not all research is positive. Secondly, the predominance of audit examination and regression in the draft framework is not surprising, since these techniques have been typical of practice where APs are required. However, this popularity does not necessarily indicate that these are the most effective and efficient methods, given the modern business environment and big data. Researchers should look beyond these more frequently used methods to other methods. Thirdly, the EAA framework proposals in Figure 18 should be explored and expanded by interested researchers and regulators. A detailed listing of the papers discussing or testing certain techniques in each audit phase is available in Table 21 of Appendix A and Table 23 of Appendix B.

It is hoped that this literature synthesis assists in the evolution of an answer for the current dilemmas facing academia, regulators, and the profession. Identification and organization of the literature will enable interested researchers and practitioners to quickly grasp the enormity of the extant research and its scope. Academia has already conducted extensive research regarding the use of analytics in the external audit and even more is required. The application of these papers towards an EAA framework maintains their relevance in the modern economy and in the modern data-driven audit.

CHAPTER THREE: BIG DATA AND ANALYTICS IN THE MODERN AUDIT ENGAGEMENT: RESEARCH NEEDS

3.1 Introductory Discussion of Business Analytics and Audit Analytics

Many different analytics terms are mentioned in these chapters, but care should be exercised when discussing analytical procedures and business analytics (BA) in the public audit engagement context because the two terms might not be completely interchangeable. Analytical procedures, according to AS 2305 (PCAOB, AS 2305 2016), are an important part of the audit process and mainly consists of an analysis of financial information made by a study of believable or plausible relationships among both financial and non-financial data. These analytical procedures could be a basic as scanning (viewing the data for abnormal events or items for further examination) to more complex approaches (not clarified by the standards, except that the approach should enable the auditor to appropriately develop an expectation and subsequently examine these expectations to the reported results).

Business Analytics (BA) that is utilized by client management and their accountants has been defined as "*the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their operations, and make better, fact-based decisions*" (Davenport and Harris 2007). BA may be further conceptualized with the three dimensions of Domain, Orientation, and Technique as shown in Table Two (Holsapple et al 2014). Often the dimension of Orientation is separated into three levels:

descriptive, predictive, and prescriptive, see Table Three (Holsapple et al 2014; Davenport and Kim 2013; Evans 2012).

The focus or context of BA for management would be somewhat different from that of the auditor. Management accountants are seeking to extract and develop insightful knowledge to enhance efficiency and effectiveness of operations, in addition to providing forecasts to enhance management decision-making. Internal auditors are seeking to verify the effectiveness and accuracy of this information. External auditors are concerned with BA as they relate to verification of the veracity of the financial statements. However, both audit tasks involve generating expectation models as well as confirmatory models. Since auditors examine business financial data, their work is affected by business analytics.

Techniques are the analytical approaches that can be described as either descriptive, predictive, or prescriptive, depending on the task of the analysis and the type of data. The more forward looking the task and the more varied and voluminous the data (big data), the more likely the analysis will be prescriptive or at the very least, predictive. Advanced or more complex BA may be defined as *"Any solution that supports the identification of meaningful patterns and correlations among variables in complex, structured and unstructured, historical, and potential future data sets for the purposes of predicting future events and assessing the attractiveness of various courses of action. Advanced analytics typically incorporate such functionality as data mining, descriptive modeling, econometrics, forecasting, operations research, optimization, predictive modeling, simulation, statistics, and text analysis"* (Kobelius 2010).

If audit clients are utilizing these more advanced BA techniques operation wide, is the auditor conducting an effective and efficient engagement by utilizing ratio and trend analysis and scanning, which are the techniques typically used and with which the auditor is comfortable (Glover et al. 2014)? When would the auditor rely more on analytical procedures over substantive detailed testing? Or, is there room in the current understanding and regulations of analytical procedures for these more complex approaches? Can analytical procedures be regarded as Audit Data Analytics?

Stewart (2015) defines: "Audit Data Analytics (ADA) is the analysis of data underlying financial statements, together with related financial or non-financial information, for the purpose of identifying potential misstatements or risks of material misstatement." This definition is illustrated by linking analytical procedures with traditional data procedures (Figure 19). ADA encompasses both the traditional file interrogation with which auditors are quite familiar as well as analytical procedures and analytics, some of which auditors may be less acquainted with. Both may be more easily understood by obtaining an understanding of the modes of ADA. Traditional file interrogation and analytical procedures are subsets of the larger field of ADA. If ADA is understood as exploratory or confirmatory in task, this task oriented approach "allows" the auditor to utilize other techniques.

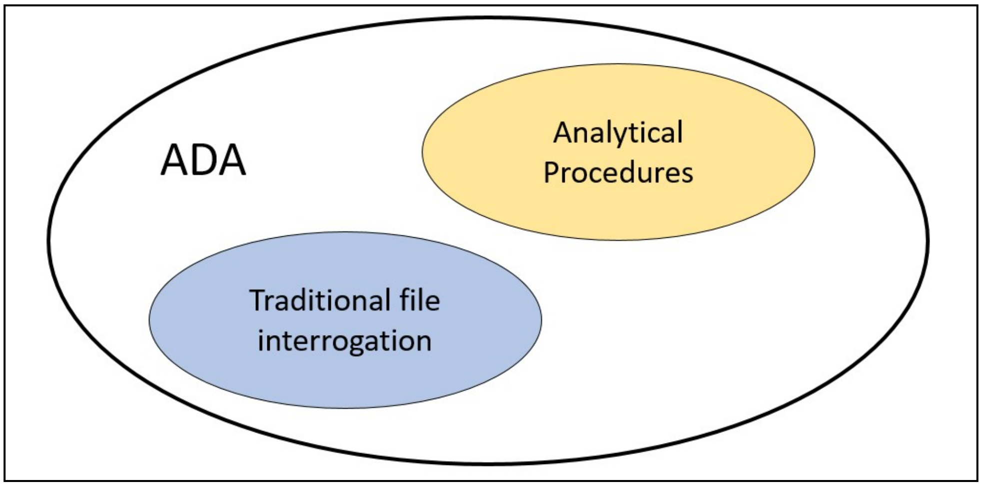

Figure 19: Linking Analytical Procedures to traditional file interrogation (Stewart, 2015)

Liu (2014) has proposed the use of Exploratory Data Analysis (EDA) (Tukey 1977, 1980) in the audit process to generate more directed and risk sensitive audit assertions for their ensuing usage through Confirmatory Data Analysis (CDA). Furthermore, Liu (2014) examined where these applications could be used in the audit process as well as their placement in extant audit standards (see Appendix A). Liu (2014) and Stewart (2015) placed EDA and CDA into the context of audit data analytics and argued for its usage as parts of audit standards. To this definition Stewart (2015) and Liu (2014) add that ADA can be exploratory and confirmatory and illustrate its functionalities. Although new or more complex methods can be proposed and even adopted by firms, it does not mean that these methods are actually promoted by the standards – instead, these new methods are simply not precluded. For instance, while regression was incorporated in the Deloitte, Haskins and Sells methodology (Stringer and Stewart 1966), its use today is not yet widely nor clearly accepted.

In summary, the standards define the task for analytical procedures in each of the three phases, but are non-committal about which techniques auditors should undertake to achieve these objectives. Whether an auditor employs more complex BA or "traditional analytical procedure" techniques seems to depend on the auditor's own knowledge of analytics and less so on the standards. The standards only provide guidance on when the auditor must use analytics, leaving the type of approach open to auditor judgement and preference. And as mentioned earlier, the auditor appetite for more complex analytics seems to have weakened since the passage of SOX in 2002. It has been proposed that any adoption by the external audit profession of either advanced analytics or big data would be due to market or business forces exogenous to the firms (Alles 2015). The recent revival of interest in ADA by the firms may be due to these forces.

This brief discussion of BA in contrast to the analytical procedures utilized by auditors in engagements provides many areas for future debate and research. Additional topics that were identified in the Introduction section are broadly summarized with these six concerns that follow.

3.2 Six Concerns Relative to Advanced Analytics in the Modern Engagement

US audit practices, methods, and regulations have evolved over the last 100 years with the constraints of auditor capabilities and the cost benefit considerations of existing business processes and technologies. These constraints stand in contrast to the new and evolving business information systems environment. The advent of computers, large storage systems, and integrated software has transformed business processes in the first wave of the information age. Their availability has brought to the front the potential of a large number of analytic methods progressively being used in business but still emerging in the external audit domain. The six next research questions identified in the Introduction chapter are discussed in detail here.

3.2.1 Should New Analytics Be Used in the Audit Process?

Perhaps this research question could be rephrased as: Should auditors expand their use of analytical procedures beyond that of scanning, ratio and time series analysis, and detailed examination? Are these techniques effective and efficient in a big data context? Basically, these questions emerge and are summarized in Table 7: Should there be more guidance regarding analytic methods in the audit? Do we know enough about these methods that this guidance can be issued? What are the tradeoffs between 100% population tests, sampling, and ad hoc analytics? The standards (PCAOB 2010, AS 1105) suggest that 100% testing would only apply in certain situations, such as: the population consists of a small number of high value elements; the audit procedure that is designed to respond to a significant risk and other means of testing do not provide sufficient evidence; and finally, the audit procedure can be automated effectively and applied to the entire population. The last condition is noteworthy, as current technologies can support automation of basic audit tests such as three-way matching and sampling, in addition to handling fairly large data sets.

The strong emphasis on judgment that exists in auditing is justified by the enormous variety of situations that complex businesses, different industries, international locations, and data structures present to the engagement team, limiting their ability to narrowly pre-set audit rules. Do modern statistical and machine learning methodologies make it possible to automate pre-set rules in many situations in order to perform procedures, derive results, and integrate these in a larger judgment? Can audit findings and judgments be disclosed in more disaggregate manner with the usage of drill-down technologies where the opinion would be rendered and broken down into sub-opinions and quantified in terms of probabilistic estimates (Chesley 1977, 1978)? Can the above be stated in terms of rules implementable in automated audit systems to continuously monitor and drive Audit by Exception (ABE) (Vasarhelyi and Halper 1991)?

Issue of New Analytics in the Audit	**Recommendations**
Should there be more guidance in the standards regarding analytical methods?	This issue should be debated amongst practitioners, academics, and regulators. Perhaps the PCAOB should open commentary.
Do we know enough about these BA methods to issue guidance?	More careful research should be conducted about which methods would be more appropriate for the assertion and audit task before guidance can be issued.
What are the trade-offs between 100% population tests, sampling, and ad hoc analytics?	This issue is discussed in depth and recommendations provided later in this paper. Also see Brown-Liburd et el (2015).
Does analytics allow for automation of many judgment oriented audit procedures?	More experimental research is needed to evaluate the possibility of automation of many judgment oriented audit processes.
Can the audit opinion be disclosed in a more quantified and probabilistic manner?	This issue is discussed in depth and recommendations provided later in this paper.
Can the above be stated in terms of rules implementable in automated audit systems to continuously monitor and drive audit by exception (ABE)?	A framework for an automated ABE system should be proposed which takes advantage of the

	big data processing and business analytics capacities of modern enterprise systems.

Table 7: Summary of the Issues regarding New Analytics in the Audit and Recommendations for Future Research

3.2.2 Which of These Methods are the Most Promising?

The literature on Big Data and Analytics methods applied to business is very large. These methods suggest different staging of the audit (audit re-modularization), changed organization (separate analytic function), changed sequencing, changed tasks, changed timing (continuous, agent driven, exception driven) (Vasarhelyi and Halper 1991) and changed personnel (more literate in IT and data; specialized) making it difficult to evaluate the literature in the context of the external audit. Appelbaum, Kogan, and Vasarhelyi (2016) have recently organized, examined and categorized this body of external audit literature. That study covers more than 300 papers published since the mid-1950's that discuss analytics in at least one phase of the audit. Due to the standards requiring analytical procedures in both the planning and review stages, these two phases are the predominant focus in the literature as is substantive testing and sampling (Appelbaum et al. 2016). Many different analytical techniques are utilized at all phases of the audit, but in an inconsistent manner. Methods that are most promising are categorized as follows:

1) Audit Examinations: transaction tests, ratio analysis, sampling, confirmations, re-performance, CAATS automation;

2) Unsupervised[1]: Clustering, Text Mining, Visualizations, and Process Mining (discovery models);
3) Supervised[2]: Process Mining (process optimization), SVM, ANN, Genetic Algorithms, Expert Systems, Decision Aids, Bagging, Boosting, C4.5 classifiers, Bayesian Theory, Bayesian Belief Networks, Dempster-Shafer Theory Models, Probability theory models;
4) Regression: Logistic, Linear, Time Series, ARIMA, Univariate, Multivariate;
5) Other Statistics: Multi-Criteria Decision Aid, Benford's Law. Descriptive Statistics, Structural Models, AHP, Spearman Rank Correlation Measurements, Hypothesis Evaluations, and Monte Carlo Study/Simulation.

These analytical models range from very simple substantive tests and routines to more complex and predictive techniques requiring significant auditor judgement. The auditor will need to determine what type of analysis gives the best quality and most efficient audit, given the audit task, the assessed audit risk, and the available data. Ideally, the auditor should be able to perform most if not all procedures to more exacting standards in a big data and continuous auditing or monitoring environment using a variety of analytical approaches. Using targeted techniques, auditors would spend less time navigating through insufficient samples and instead, identify and almost immediately examine the transactions of high risk.

---

[1] Unsupervised approaches are those techniques that draw inferences from unlabeled or unknown datasets since there is minimal hypothesis of the results based on labeled responses

[2] Supervised approaches are those techniques that draw inferences from labeled or known dataset types, otherwise known as training data

Auditors selecting these more complex techniques need to understand them in terms of their benefits and limitations. Furthermore, the tasks of risk assessment, substantive procedures and tests of controls may be different when 100% of the data is examined (Yoon 2016). For example, if auditors are examining 100% of items in the population (PCAOB 2010, AS No. #1105.24), the emphasis and reason for testing internal controls should change. Internal Control testing has been prescribed in the regulations (American Institute of Certified Public Accountants [AICPA] 1997, SAS No. #80) to supplement substantive testing for validating sampling results when auditors have limited access to data. It has been suggested (IAAE 2016 p. 18) that internal controls testing in an Audit by Exception type of environment could provide some assurance regarding data quality.

To summarize the issues of which methods are the most promising (Table 8) given the audit task as defined by the standards: A new environment of assurance is emerging where automation of controls, full population testing, and analytic methods will interplay. Research is needed on modern analytic methods to establish: their applicability in different instances, their cumulative effect, their ability to be formalized, their classification (creation of taxonomies of analytic methods and data structures[3], and their quantification.

A set of questions arises with the application of analytics that must be tested in the field. Would a safe harbor experimentation (a la XBRL) process be needed for the testing of approaches? Although in the traditional environment a yes/maybe/no

[3] The AICPA has created the Audit Data Standard (Zhang et al. 2012) to guide in the formalization of data to be received in the audit, its classification (into cycles), and its measurement.

attestation is provided, the new proposal provides information of audit results in at least five areas where needed. How would these results be disclosed?

Issue regarding Which Methods are most promising	**Recommendations**
In what circumstances would modern analytical or more complex methods be appropriate?	Research should examine if the current standards regarding sampling, selection of specific items, or 100% tests could be expanded.
What would be the effect on the engagement, the firm, the standards?	This question could be incorporated in the same research above.
Could these approaches be formalized, if not industry wide at least internal to the firm?	This question could be incorporated in the same research question above.
Who would classify or standardize these approaches (create a taxonomy of methods and data structures for defined audit tasks)?	Perhaps this process could evolve under the guidance of the AICPA in collaboration with academics and practitioners.
How would these approaches be quantified?	A quantification framework could be proposed and demonstrated.
How would these approaches be tested in the field? Sand box approaches accompanied with successive levels of adoption? Would these be provided a safe harbor?	This could be part of the AICPA initiative with firm support (perhaps a RADAR project) and academic input.
Again, how would this affect the audit opinions? Could these modern analytical methods facilitate more transparent and quantitative disclosure?	A framework or guidance for a more detailed and quantitative opinion disclosure should be developed and proposed.

Table 8: Summary of issues regarding which methods are most promising

3.2.3 Where in the Audit Are These Applicable?

The traditional organization and processes of the audit as defined in the current standards will be affected in many ways by the emerging environment and its disruptive technologies. If some form of Audit by Exception (ABE) (Vasarhelyi and Halper 1991) emerges whereby the audit process is activated by alarms triggered in data streams, and a plethora of new analytics emerge, clearly the sequence of events will be transformed and the applicability of analytic methods expanded. Furthermore, there will be ubiquitous use of techniques such as visualization, and multi-complementary use of many analytic methods. Visualizations are used heavily in business management to explain the results of analysis (Dilla et al. 2010; Kohavi et al, 2004). Many techniques exhibit varying strengths and weaknesses and are more beneficial when applied in combination rather than separately. The sequencing (or simultaneity) of events will change as automated use of data analytics will precede / or coincide with the more traditional audit examination which may progressively be reduced. For example, today the audit engagement typically progresses as shown in Figure Two but is envisioned to eventually innovate to a more Audit By Exception (ABE) approach (Figure 20).

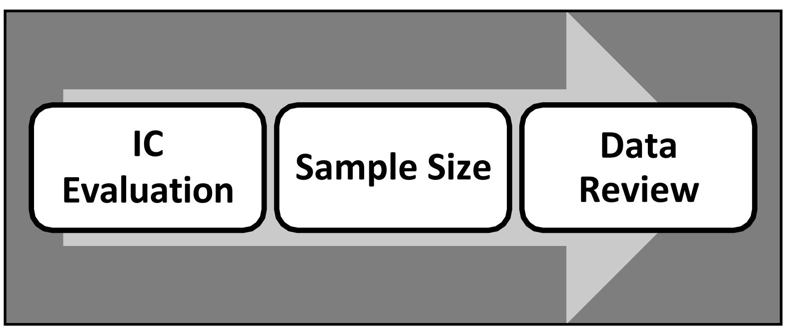

Figure 20: The current typical audit plan

The above process, which drives most current engagements, is sample driven; in a more data driven environment the examination process would be analytically reviewed, audited automatically, and exceptions or outliers would be subsequently examined in detail (Figure 21).

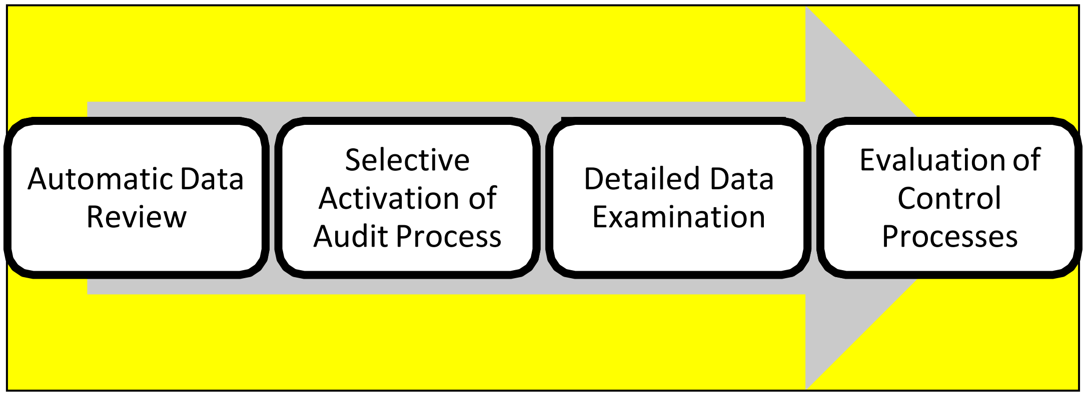

Figure 21: In a more Data Driven Process, Audit by Exception (ABE) of audit examination

However, in this ABE approach the auditors may face a different challenge: testing all of the transactions may produce thousands of exceptions (Dohrer, McCullough, and Vasarhelyi 2015) if the threshold definition of a material deviation is set too high. That is, the threshold approach for sampling most likely will not work in ABE; the threshold should be more precise to eliminate the "false positive" exceptions. The standards require that all exceptions should be examined (PCAOB 2010, AS No. #2305, AS No. #2315), but this was mandated for sampling (IAAE 2016 p. 17). In an ABE context, if the tests were not configured correctly, there could be an unreasonable number of exceptions to investigate as required. Some auditors have performed additional tests to "explain away" many of exceptions and categorize the resulting few as "Exceptional Exceptions" (Issa et al. 2016). Clearly auditors will need to possess a broad and comprehensive knowledge of analytical techniques in an ABE environment.

The level of automation of the audit, and as discussed before, the availability and comfort with analytical techniques, the competences of the auditor, and the circumstances and assertions of the specific audit process will guide the locus of the application. As such, ABE is a more advanced audit approach, reflecting the confluence of automation, advanced analytics, and revised regulations. Issues that may emerge during this process could be as follows (Table 9): How different are the objectives of Internal Audit and External Audit in the current context (Li et al. 2016)? Isn't there a substantive overlap between business monitoring and real time assurance?

Considering that there is substantive overlap in data analytic needs, are the traditional three lines of defense (Freeman 2015; Chambers 2014) still relevant[4]? Traditional auditing has a retrospective approach, as traditional technologies did not allow for other approaches - can the current environment allow for a prospective look and to what extent? What parts / procedures of the audit are fully or partially automatable? Will the disruptive changes (Christensen 2013) be allowed by the leading audit firms?

Can the key contingencies in the audit be formalized? In the same line, but extending expanded testing and reporting, should quantitative guidelines be issued for ABE and its structures, and should within period results be disclosed as part of the auditor's report? The succinctness of the traditional report is not necessary any more, and drill downs on the results of Critical Audit Matters (CAM) examination, their details, and other information is possible.

Issues about where in the audit these analytics would be applicable	**Recommendations**

[4] There should be effective risk management functions within a company. These monitoring and assurance functions have been modeled as the "Three Lines of Defense" by the IIA. This model serves as an example, where: 1) the first line of defense represents functions that own or manage the risk; 2) the second line of defense, where there are functions that specialize in risk management and compliance; and 3) the third line of defense, where there are functions that provide assurance

How would the objectives of internal and external audit differ in this context?	Research should examine the areas of convergence and separation in the context of integrated enterprise systems, analytics, and big data
Isn't there a substantive overlap between business monitoring and real time assurance?	This has been alluded to in earlier research but should be re-examined if the assurance process changes
Considering that there is an overlap in data analytic needs between different functions, how relevant are the three lines of defense?	Recent works by COSO have questioned the feasibility of the three lines of defense – however, the independence of assurance must be maintained, which is an area for future research. There are many possibilities for the three lines of defense.
What parts of the audit engagement are fully or partially automatable? Would auditor judgment eventually be replaced with prescriptive analytical algorithms?	This area could be examined at depth with varying levels and moments of audit automation, factoring such variables as judgement and interim testing
Would leading audit firms allow such disruptive changes in engagement practice, absent regulation changes?	Would these firms be willing to be key innovators in the assurance side? (Perhaps if they were to be allowed a sandbox or safe harbor?)
Can the key contingencies in the audit be formalized?	These should be examined and articulated with frameworks/guidelines embedded in an expert system
If the annual audit opinion can become more informative, as per recent CAM reviews, why stop there? Why not issue CAM level quarterly reports and reports on demand?	The recommendations regarding this issue are discussed later in this paper. CAM reviews could serve as the foundation of a more quantitative opinion report. Other possibilities evolve for an immutable real-time seal of the data and its assurance

Table 9: Issues regarding where in the audit these methods would be applicable

3.2.4 Should Auditing Standards Be Changed to Allow / Facilitate These Methods?

In general, the aforementioned meetings between the AICPA's ASB and the ASEC committee have concluded that the standards do not forbid the usage of analytics, but it can be argued that the standards, and the economics of external audit, make analytics more difficult or in some instances impractical if not nearly impossible to use. The lack of a more detailed discussion of appropriate analytical techniques within the standards, when placed in the context of a highly competitive business environment, does not encourage the profession to explore new techniques even in the face of big data and automation. The use of more automation and analytics in the engagement, particularly in a big data environment, generates these additional issues (Table 10):

- The economics of the audit is encumbered by a series of anachronistic requirements still being enforced by the PCAOB. Consequently, the pricing of the audit, in a competitive environment, leaves little space for additional analytics even if these give stronger assurance of fair representation. Furthermore, what would be the cost versus benefit trade-off with the usage of analytics? Or, would there be a point where the cost of conducting a sample driven audit exceeds that of ABE audit? When would the additional assurance derived from the analytic results justify the cost of their application? Even further, if a particular analytics method is more powerful and uncovers issues that were not previously detected, what would be the liability of the accounting firm, particularly if these issues were also present in the prior years? (Krahel and Titera 2015, p. 418)
- Sampling requires laborious follow ups on abnormalities detected, but in a population of millions or hundreds of thousands there is little to be gained from picking 25 transactions and reviewing them (Dohrer, McCullough, and

Vasarhelyi 2015). Do any areas of the modern audit exist where these small judgmental samples still make sense (Elder at al. 2013)? In juxtaposition to the current requirements, would the auditor then need to justify the use of sampling in circumstances where 100% of the data would be available for testing?

The audit research literature itself has been scant regarding auditors' sampling decisions in the context of economic and competitive pressures, regulations about statistical sampling, as well as how to effectively extract meaningful results from the sampling (Elder et al. 2013, p. 103). Auditing standards (PCAOB 2010, AS No. #2315) define sampling as "the application of an audit procedure to less than 100% of the items in an account balance or class of transactions for the purpose of evaluating some characteristic of the balance or class.", The auditor may choose to select all items for testing if the level of sample risk from possible erroneous decisions is too high (AS No. #2315.07).

There is little guidance as to when 100 percent testing would be more appropriate than selecting specific items. In the standards about Audit Evidence (PCAOB 2010, AS No. #1105.22-.29), sampling is not recommended when the data population is small and/or not homogeneous, when there appears to be significant risk, when there are key items that should be examined, when threshold tests should be applied, nor is it suggested when audit procedures can be automated effectively and applied to the whole population. In the standards regarding sampling (PCAOB 2010, AS No.

#2315.07), the auditor should weigh the cost and time to examine all of the data versus the perceived degree of uncertainty from sampling and non-sampling risks, and judge accordingly. Consequently, the practice of sampling has become embedded in basic public auditing practice. PCAOB examinations have been very strict favoring sampling against analytical methods.

- Furthermore, Elder et al. (2013) were unaware of any literature that addresses the auditor's decision to use audit sampling of any type (Elder et al. 2013, p. 111) and suggested that future research should address the issues of when sampling would be appropriate and when other types of tests would negate the need for sampling. In response, Yoon (2016) discussed how substantive analytical procedures (SAPs) applied to 100 percent of the data (with the use of computer assisted auditing techniques) could potentially provide a more efficient and effective audit evidence than sampling, particularly in a big data environment. Perhaps for audit engagements where the client is collecting or analyzing all of the transactions and the auditor is using automated audit software, the standards could more clearly establish that 100 percent tests using substantive analytical procedures would provide efficient, sufficient, and appropriate audit evidence.

For example, three way matches used to be performed manually and reviewed manually. Now advanced accounting systems and ERPs perform these automatically. Is this performance audit evidence, new analytics, or just automation? If considered automation, how do the audit standards take this into consideration? Is there a difference between automation and analytic methods? (Dohrer, McCullough, and Vasarhelyi 2015) If such automation is viewed as preventive internal control, then how does it change the balance between control testing and substantive testing in auditing the modern highly automated enterprise environments?

- In highly automated accounting systems many analytics or pre-programmed apps will depend on some form of "audit data standard" (Zhang et al. 2012)[5]. These apps will run frequently or constantly (Vasarhelyi and Hoitash 2005). This form of evidence may use external and internal data (Brown-Liburd and Vasarhelyi, 2015) potentially from external sources like social media, thus providing valuable tertiary audit evidence that may be used to complement / replace current tests. Would these need new guidance? Are the current guidelines for traditional audit evidence the same for external or internal big data, particularly social media? What qualities should these data possess in order to provide reliable audit evidence?
- It has been shown (see e.g., Hoitash et al. 2006) that the performance of audit analytics is significantly improved if the models incorporate contemporaneous peer company data. Conceivably, contemporaneous peer company data should be considered as legitimate sources of information for obtaining an understanding of the relevant industry and the client's position, as outlined in the standards for risk assessment and review (PCAOB 2010, AS No. #2110, AS No. #2810). Large public accounting firms typically audit multiple peers in

[5] The AICPA has published online a series of voluntary suggested audit data standards: http://www.aicpa.org/InterestAreas/FRC/AssuranceAdvisoryServices/Pages/AuditDataStandardWorkingGroup.aspx

the same industry, and they could create large internal data warehouses to share such data among the engagement teams during the audit. The current strict interpretation of audit client confidentiality rules causes the firms to err on the side of caution and disallow any sharing of data even though such data would never leave the confines of the firms. New guidance interpreting client data confidentiality as being safeguarded within a firm (and not within an engagement team) and specifically allowing audit client data sharing among different engagement teams would greatly enhance the performance of audit.

Should the standards change to facilitate these methods?	**Recommendations**
What would be the cost versus benefit trade-off with the usage of analytics in the current regulatory environment?	This issue should be examined as the cost benefit of more advanced analytics may be a major variable affecting the use by firms
What would be the breaking point of sample driven audits versus 100% tests resulting in ABE?	The effectiveness and efficiency of the two audit approaches should be examined in future research. This issue has been conceptually addressed in Yoon (2016)
When would the value derived from the additional assurance provided by analytical results justify their incremental cost?	Collaborative research efforts between academics and firms would be appropriate to address this issue
If more powerful analytics uncovers issues that were not previously detected, what would be the liability of the audit firm, particularly if these issues have been on-going?	This is an issue that the regulators should address, with input from the firms and researchers. This may relate also to earlier “safe harbor” questions
If the auditor has access and ability to test 100% of the dataset, would there still be justification for the use of sampling?	This is an issue that research should address, allowing for time, accuracy, and cost calculations for sampling versus 100% tests
Is there a way to quantify the evaluation of the cost and time to run 100% tests versus the perceived liability of sampling risk and judge accordingly?	This is an issue that the regulators should address as part of the preceding question
Are 100% tests new type of audit evidence or just automation?	This question could be examined along with other issues relevant to big data

If these tests are considered automation, how do the standards take this into consideration? Should the current solution of greater reliance on internal controls be quantified?	This is an issue that the regulators should address, with input from the firms and researchers. The controls testing and verification process as it relates to an IT audit and the reliability of information generated within a system may need clarification/quantification.
Is there a difference between automation and analytic methods? Isn't automation basically the automated application of analytics?	This is an issue to be considered in future research efforts by academics, as part of a scoring framework for audit evidence
If such an automation is viewed as a preventative internal control, then how does it change the balance between control testing in auditing the modern highly automated enterprise system?	This is an issue that the regulators should address, with input from the firms and researchers.
Would evidence from external sources such as social media require new guidance?	This question should be examined in detail given the veracity issue with external big data. Guidelines regarding normative expectations should be established – this evidence should be scored as part of the quantitative evidence framework
What qualities should this data possess in order to provide reliable audit evidence?	This query can follow the recommendations proposed previously in the big data external evidence guidance discussion
Could the standards allow firm industry knowledge to be supplemented with anonymized confidential peer company data?	This is an issue that the regulators should address, with input from the firms and researchers.
Could new guidance be offered that defines client confidentiality as being firm wide in scope and not limited to an engagement team?	This is an issue that the regulators should address, with input from the firms and researchers.

Table 10: Where should the standards be changed to allow/facilitate these methods?

3.2.5 Should the Audit Report Be More Informative?

PCAOB Release No. 2016-003 proposes, concerning an unqualified opinion, that the audit report disclose "Critical Audit Matters" (if any) in areas such as estimates, audit judgments, areas of special risk, unusual transactions, and other significant changes in the financial statements. This proposal [6]poses a series of interesting questions worthwhile of research (Table 11): Is the level of proposed disclosure adequate in terms of quantification of these critical audit matters or is it falling back into the comfort zone of the traditional auditor? After all, substantive industry resistance was found to the initial proposal (PCAOB, 2013[7]). Would some of these Critical Audit Matters (CAMs) provide disclosures that are more disaggregate, or more informative than the traditional audit reports?

Could there be preferable schemata of quantification, or quantitative guidelines for estimates, audit judgments, areas of special risk, unusual transactions, or other significant changes in the financial statements? Should these schemata be determined by the standard setter? On a longer range, if the auditor is using/ relying on ABE should there be a real time seal or similar device that would allow investors to know on an immediate basis that auditors are monitoring systems and they seem to be doing well[8]?

Should the audit report be more informative?	**Recommendations**
Is the level of disclosure appropriate for more advanced analytics and quantification of critical audit matters (CAMs)?	A framework for appropriate disclosure should be developed

[6] See also Lynne Turner's comments (https://pcaobus.org//Rulemaking/Docket034/ps_Turner.pdf).

[7] PCAOB Release No. 2013-005, August 13, 2013, Docket Matter No. 034, The Auditor's Report on an audit of Financial Statements When the Auditor expresses an Unqualified Opinion. This report discusses the auditor's responsibilities regarding certain other information in certain documents containing audited financial statements and the related auditor's reports and related amendments to the PCAOB standards.

[8] This type of continuous assurance would work better with some form of more frequent/ continuous reporting.

Would some of these CAMs provide disclosures that are more disaggregate or more informative than the traditional audit reports?	This is an issue that researchers and regulators should examine as a more informative CAM component of the audit opinion is formulated
Should there be quantitative guidelines for estimates, audit judgments, areas of special risk, unusual transactions, or other significant changes to the financial statements, and if so, by whom? Regulators? Researchers?	This is an issue that the regulators should address, with input from the firms and researchers.
Or projecting in the future, if the auditor is relying on an ABE assurance protocol, why shouldn't audit reports be generated more frequently or on a just-in-time/on demand basis?	This could be one aspect of a forward-looking paper by academics that conceptualizes a grand vision of the future public audit. This could be a new form of service by auditors that probably now is forbidden by SOX.

Table 11: Should the audit report be more informative of Critical Audit Matters (CAMs)?

3.2.6 What are the Competencies Needed by Auditors in This Environment?

As mentioned above, the application of analytics in the external audit is attracting substantial attention from practice and academia. EY[9] and the AAA[10] among several others have brought together these two groups for constructive dialogues. Auditor education and familiarity with analytics has been positioned by the standards as a limiting factor regarding which techniques to apply in the engagement (PCAOB 2010, AS No. #2305). Papers such as Tschakert, Kokina, Kozlowski, and Vasarhelyi (2016) and Appelbaum, Schowalter, Sun, and Vasarhelyi (2015) have discussed the issues facing audit education. In general, some conclusions could be drawn:

- Accounting faculties tend not to be prepared to teach analytics.

[9] EYARC 2015, June 17/18 2015, Dallas Texas.

[10] AAA, Accounting is Big Data, September ¾ 2015, New York, New York.

- There is a widespread feeling that students are not receptive to learning analytics.
- The accounting curriculum is too full to add more IT, statistics, and modeling.
- As the CPA exam does not include these topics, there is little motivation by students for their addition to the curriculum of study.
- Firms will tend / or already have hired specialist groups from non-accounting backgrounds. These groups, as in IT audits (Vasarhelyi and Romero 2014) will be external to the audit team and brought in if the manager of the engagement setting up the audit plan sees fit.
- Practitioners are also not prepared and their internal audit practices have not caught up properly with these issues.
- Firms have been developing software to improve their processes but feel curtailed by the PCAOB examination process.

These factors lead to a series of educational research questions and potential projects that are paradigm changing (Table 12): If the curriculum is too full, if memorization in the age of google is a different consideration, and if the domain of coverage is too large, then what educational structures and what types of certificates should be used /developed?

Should the CPA profession expand competencies or progressively rely more and more on specialists from other domains, potentially using other (non CPA firms) to provide these competencies? Should the set of CPE requirements of the profession be reformulated in terms of a life-long-learning approach where new required skills are defined and progressively required in the accountants learning/ competency profile? Who should manage this learning profile, and who should set the requirements? Should there be a much wider set of accounting specializations with coordinated competencies? Should there be quantification of the different types of accountant skills? And some of these acquired through on the job activities and related experience?

Issues about auditor competencies	**Recommendations**
In this day of Google and other IT tools, should the curriculum be filled with rote memorization tasks?	This topic should be examined and developed by academics with guidance from the AICPA
What types of education requirements, structures, and certification should be developed?	This topic should be examined and developed by academics with guidance from the AICPA
Should the audit profession move more towards the use of IT and analytics specialists in the engagement or is there room for this additional knowledge?	This topic should be examined by practitioners and academics in a behavioral study setting
Should the CPE requirements of the profession be reformulated to reflect these new learning skills/requirements?	This topic should be examined and developed by academics with guidance from the AICPA
Should there be a much wider set of accounting specializations with coordinated competencies?	This topic should be examined and developed by academics with guidance from the AICPA
Should there be quantification of the different types of auditing skills?	This topic should be examined and developed by academics with guidance from the AICPA

Table 12: What are the competencies needed by auditors in this environment?

3.3 Technology Adoption Issue: Evolution Towards a New Audit Environment of Big Data and Audit Analytics

"It has also been shown that many internal audit procedures can be automated, thus saving costs, allowing for more frequent audits and freeing up the audit staff for tasks that require human judgment." (AICPA, 2015)

It has been proposed in other technology adoption settings that such automation changes are best considered as *evolutionary* instead of *revolutionary* (Kuhn and Sutton 2010). The topics and suggestions mentioned in this paper may seem extensive in scope and massive in undertaking. These issues could serve as either motivators or impediments to the use of big data and audit data analytics (BD/ADA) by the external audit profession.

Ideally, it would seem that the goal for BD/ADA adoption by the profession would be to save costs and attain greater efficiencies and effectiveness in the audit process. However, it is conceivable that impediments exist that would dampen enthusiasm for BD/ADA adoption and these conflicts may be similar to those of other technology initiatives. Here are just a few of the issues that are proposed as being relevant to BD/ADA adoption (Table 13):

Issues for BD/ADA adoption	**Recommendations**
What are the goals/benefits/costs for each stakeholder/involved party?	Key drivers and motivating factors should be identified by firms, regulators, and clients. These should be discussed in terms of cost benefit analysis and effectiveness
Who should be the champions for this change?	To what degree and when would auditors use BD/ADA and who decides this? Who would be the main champions for this change?
How would this process develop?	To what degree and when would auditors use BD/ADA and who decides this? Should current

	audit procedures and regulations be changed prior to use of BD/ADA?
Who measures the effectiveness of using BD/ADA vs. not using and by what metrics?	Effectiveness and cost benefit analysis evaluation results may differ between stakeholders. Process of measurement metrics and expectations should be developed.
How would BD/ADA adoption take place at the firm level and regulatory level?	This question ties in with the process development (third) question
Would audit procedures need to be re-aligned to fit this new engagement environment?	Should current audit procedures and regulations be changed prior to use of BD/ADA?
How would auditors best prepare for these tasks that require more judgment and less routine work?	How would firms and regulators go about best preparing practitioners to transition to more judgement based and analytical approaches?

Table 13: Issues that might impact BD/ADA adoption

The literature regarding technology adoption is huge in the audit, accounting, and AIS disciplines. This paper does not attempt to synthesize this literature in support of this discussion; instead, a few select papers are highlighted and a very scant outline for BD/ADA adoption is suggested for future research. For instance, the Information Fusion process that Perols and Murthy (2012) propose could be applicable here in the context of BD/ADA adoption. Kuhn and Sutton (2010) present research challenges that could correspond with BD/ADA in the area of regulatory/adoption/judgment and decision making challenges. Likewise, the "messy matters of Continuous Auditing (CA) adoption" which Hardy (2015) presents may be applicable to ADA/BD.

It has been suggested (Alles et al 2008; Geerts et al 2013) that the transformation of manual processes to that of automation is best accomplished incrementally. Geerts et al (2013) and Dzuranin and Malaescu (2016) provide a framework based on Design Science for such an integration. Vasarhelyi (2013) proposes a four step process based on the work of Parasurman et al (2000). According to Parasurman et al (2000), human information processing and its evolution from man to machine can be divided into four phases: 1) information acquisition; 2) information analysis; 3) decision selection; and 4) action implementation. In the Alles et al (2008) proposal, each such successive step should be undertaken methodically once benefits from the previous steps have been realized (Figure 22).

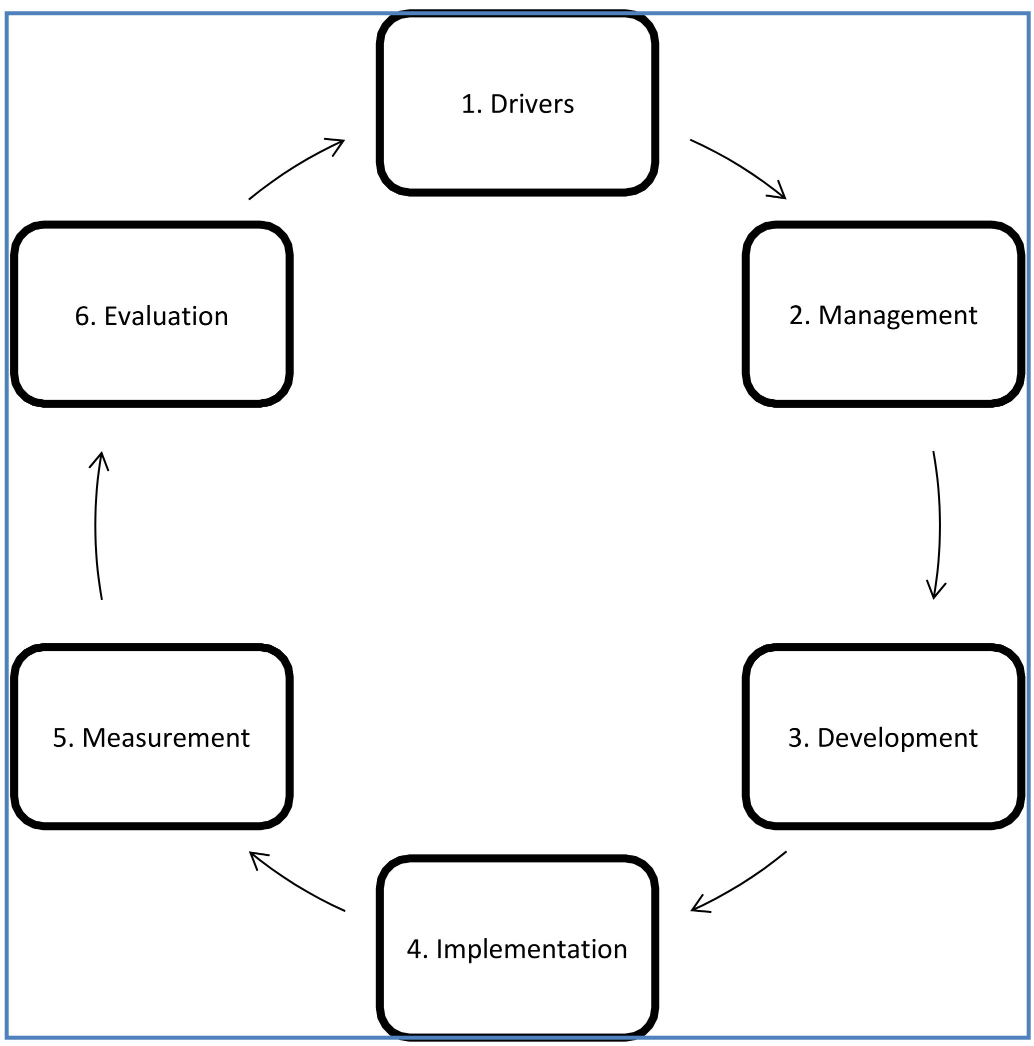

Figure 22: Different stages of one process cycle of incremental change

Furthermore, in the Alles et al. (2008) and Dzuranin and Malescu (2016) frameworks, successful change is more likely to occur if the manual process is re-engineered first to support the eventual automation. In the Alles et al. (2008) proposal, the first step of the process cycle is the consideration of the drivers of change and endorsement by management; the second step in the process is the development and the actual implementation of the components that would enable this change; the third step consists of management, or baseline measurement and evaluation of the solution. This process cycle is repeated for every level of automation transformation in an incremental fashion. Such a process cycle approach could also apply as an incremental use of analytics and big data by the public audit profession.

The initial drivers for the use of analytics and big data by external auditors are already in place, with the increasing complexity of client transactions, analytics, and data sources and the subsequent increase of audit risk to the engagement team if analytical procedures are manual and overly simplistic (Alles 2015; Bedard et al, 2008). Firms are already embracing diverse descriptive approaches (Dilla et al, 2010); it could be argued that some practitioners are about to embark on the next phase, the adoption of more predictive analytics. Basically, firms are discovering that manual and simplistic analytical procedures and data sources create an audit which is more likely than not to be inefficient and ineffective in a big data context. Many firms are investigating ways to integrate more advanced analytics in their engagements, but this initiative is progressing cautiously (Alles 2015). It is suggested that many of the research issues discussed here in this paper will need to be examined in the context of an incremental approach, as illustrated in Figure 23. Figure 23 illustrates how the process flow of Figure 22 could be integrated incrementally to incorporate advanced analytics and big data into practice.

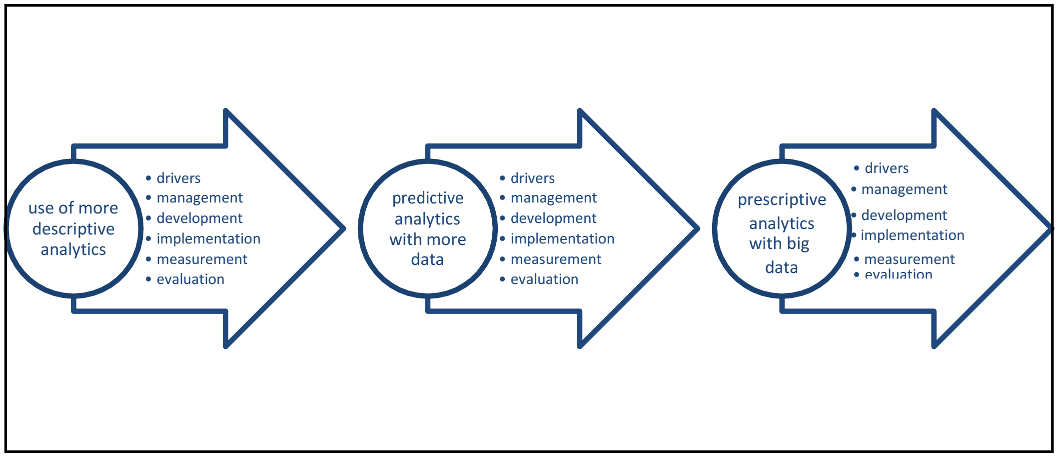

Figure 23: Three possible cycles of adoption for the use of more advanced analytics and big data by the public audit profession

This incremental approach may already be observed to some degree in the audit process – while some manual procedures have been automated, other audit procedures have not. Many audit tests may be conducted on 100% of the test population using Computer Assisted Auditing Techniques (CAATs) software packages (Wang and Cuthbertson 2015). These CAATs can perform analytics very efficiently and quickly and can interface and link easily to the client's system. Although not all CAATs software packages are equipped to handle big data, this limitation will eventually be solved. CAATs are used by auditors on many engagements for GL tests, three way matches, detail tests, and sampling for example. However, these tests do not run automatically but are manually selected by the engagement team. The auditor selects which analytical procedures or tests to run and attributes to examine in the tests of assertions for a particular audit objective.

What follows are expanded recommendations for research regarding several of the challenges mentioned earlier in the six questions.

3.4 Expanded Example of Using Advanced Analytics with CAATs for Testing of the Entire Data Population

What frequently occurs with running analytics on 100% of the population, such as commonly occurs with CAATs, and the allowable parameters for deviation are set to a low level because the transactions have a high inherent probability for error or personal use (see first research issue)? It is highly probable that too many transactions would be flagged as exceptions. These exceptions may be so numerous so as to create an alert flood (Brown-Liburd et al 2015). The flagged transactions, or those transactions tested for the assertions of accuracy and completeness in the context of client implemented controls, are too numerous for efficient and effective detailed examination. For a large dataset or big data, this exception file could have thousands of rows whereas with sampling the transactions selected for detailed examination might not even be 100 transactions.

Faced with such a voluminous exceptions set, the auditor may not know how to proceed. Testing all of these transactions in detail with corroborating evidence is too onerous in cost and time; additional analytical procedures should be conducted to prioritize these results. To serve as an effective and efficient replacement of sampling, the auditor should examine ways to reduce the size of this flagged transactions dataset. Auditor judgement is now required, as mentioned in Issues Two and Four.

At this point, engagement staff may be experiencing paralysis of choice, a state of being overwhelmed by the options available to examine this data further. The auditor could be affected by Information Overload, or the condition of simply receiving too much information (Brown-Liburd et al 2015). However, research suggests that under certain conditions such potential negative decision making effects can be addressed. For example, decision-makers with sufficient knowledge do not experience decision making paralysis (Scheibehenne et al. 2010). Also, the ease with which options can be categorized moderates the negative effects of overload (Mogilner et al. 2008). Mogilner et al. (2008) argue that categories make it easier to process the available choices and decrease the stress of making a decision, especially when the situation is unfamiliar.

One type of analytical approach, a guided expert or structured decision making system, is suggested to mitigate these information processing difficulties that the auditor is experiencing (Brown-Liburd et al 2015; Parasuraman et al 2000). One suggestion for future research would be experimental: observing any differences in auditor judgement and performance with the application of analytics to an exceptions dataset with and without a structured decision expert system. It is quite possible that firms may need to rely on the analytical knowledge of expert systems to guide inexperienced audit staff with the application of more advanced techniques to more complex datasets. It is also possible that audit technology should be integrated in a more automated fashion to facilitate auditor competencies.

3.5 Expanded example of Audit by Exception as the End Result of the Evolution of BD/ADA Use

Audit technology and methodology are highly interconnected with the way information is processed and the available capabilities as was described earlier. The evolution of available information and analytic technologies transforms how companies measure their business, how they interact with their clients, the products they produce, the method of their management and of course the way they verify (theirs' and other's) business. For example, third party probabilistic verification (Brown-Liburd and Vasarhelyi 2015) was not even conceivable before the internet and social media existed. IT audit did not precede computers, and it took years after these became commonplace in business for the field to emerge, leading to progressively disappearing paper source document verification. Similarly, as audit clients adopt more predictive and prescriptive analytics, it is hoped that these approaches will merge in the profession with audit examinations and analytical reviews, resulting in greater use of analytics and big data and eventually evolving to an Audit by Exception (ABE).

Audit methods have been retroactive (looked backwards) as they have relied on some degree of manual verification of source documents or third party verification of balances through manual confirmation. Consequently, in a big data environment, operational economics would make very onerous the manual verification of all documents, the re-performance of controls, or even analytic calculations that relied on statistics. Overall, the expected value of assurance efforts must be larger than its costs. Once manual efforts are voluminous they become very expensive. This need for operational efficiency and effectiveness in the engagement could be key drivers serving as an impetus towards the use of more analytics and big data (Alles 2015).

A series of technological events have brought in many revolutionary changes in business processing, challenges in their audit, as well as facilities for audit performance. Figure 24 below displays some of these elements. This is a symbolic representation of the environment that the public company auditor is progressively encountering in many engagements. Most notable of which is the cloud, or cloud computing, that brings together the communications ubiquity of the Internet with large storage capabilities creating organizations whose information structure is distributed but integrated, with high level of redundancy, and enormous amount of storage.

This environment is dynamic, with transactions and data streams available in close to real time. Data is flowing from many connected devices such as sensors and smart phones (Dai and Vasarhelyi 2016), from internet click stream traffic logs, and social media. This data is then almost simultaneously analyzed within the corporate information system. The client is monitoring and analyzing 100% of the available data.

The audit team should also have access to the same data in this ideal dynamic analytical model. The access could be either in tandem with the client or less ideally at a later time in a more batch oriented feed. The Audit by Exception (ABE) audit would be maximizing the potential value available for both the auditor and the client in this continuous real-time information environment. The continuous assurance process would be tied naturally to an ABE process that may lead to a different organization of the assurance process.

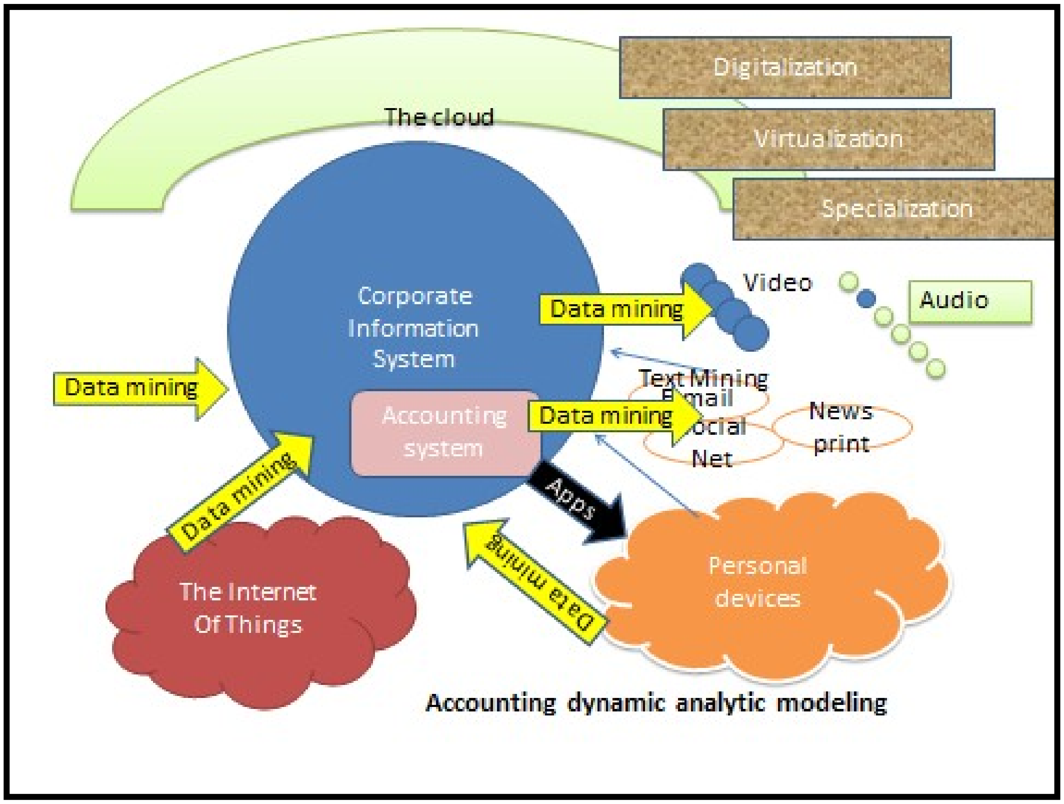

Figure 24: The evolutionary environment (adapted Liu and Vasarhelyi, 2014)

However, this dynamic model of ABE raises issues that should be mentioned and discussed: How to set the timing of performing an analytics based assurance function? Should (Vasarhelyi and Hoitash 2005; Vasarhelyi, Nelson, Kogan, Srivastava, and Lu 2000; Wooldridge and Jennings, 1995) automated applications implementing analytics be run very frequently or intermittently? The timing of these analytical tests may be regarded as similar to that of the continuous activities/interim testing phase of the audit. In this scenario, auditors may deploy analytics more frequently in a less risky client environment. ABE may occur in tandem to real time process flows or in a batch process, according to auditor judgment.

How can the sufficiency of evidence be established to protect the auditor from liability in case low priority audit data analytics (ADA) exceptions are not investigated, but are later revealed to be problematic? This liability is one that the auditor faces regarding evidence collection – the risk that an initially low priority exception becomes high priority at a later date. Traditionally, if the auditor met the requirements of the standards, there shouldn't be assigned liability. However, if 100% of the transactions have been collected and analyzed as part of the evidence collection process, such as may exist in an ABE environment, the liability of the auditor may need to be redefined. Potentially there could be less leeway for such oversights when the auditor has presumably examined all the data.

In an ABE approach, transactions that violate certain thresholds are considered to be more suspicious and are flagged for analysis. In this scenario, there are many more instances than used in sampling, and presumably the ABE system has flagged all of the abnormalities or high priority instances that could be identified. Since these instances are far too numerous for every one of them to be investigated (e.g., 25-65 from sampling versus 20,000 from ABE), the auditor must subsequently evaluate these transactions with additional analytical techniques, to identify those instances that seem to be less or more critical. This stage of the audit process in an ABE context would appear to increase the auditor's risk of liability, as it would require considerable auditor judgement. This issue was discussed in a slightly different context in the previous sub-section. The standards should be amended to provide guidelines for setting these minimum parameters based on the Inherent and Control Risks of the evidence and technique being used.

There exists great potential of a more objective quantification of the audit opinion with the application of Audit Data Analytics (ADA). A more quantitative opinion by the audit profession could contribute great social and business value. Audit opinion currently exists as a pass/fail summary, even after the audit team has conducted extensive detailed testing. The standards themselves (AS 2810) describe a rich compendium of information and additional analytical procedures in the review stage that ultimately result in this summary opinion. The levels of detail that auditors attain during the engagement can only increase with the emergence of ABE and big data. It would undoubtedly benefit the users if this highly insightful information were included in the final audit opinion.

The challenges posed by measuring the amount of evidence provided by ADA, as well as by the quantification of the concept of materiality (acceptable relative error), are formidable. Major advances in this direction will likely have the additional benefit of making it possible to quantify the audit opinion by disclosing in this opinion the actual measures achieved in the audit. An even more useful disclosure would specify the "confidence intervals" for all the accounting numbers reported. Such disclosures would make an audit opinion much more relevant by allowing various users of the audited financial statements to calibrate their reliance based on the decision tasks at hand.

3.6 Discussion and Further Research Issues

Modern audit engagements often involve examination of clients that are using big data and analytics to remain competitive and relative in today's business environment. Client systems now create and acquire big data and apply advanced analytics to generate intelligence for decision making. However, the public accounting profession is still bound by regulations that may have been applicable years ago but whose relevance should be re-examined today in this modern business environment. There are numerous issues surrounding the standards, practice, and theory of audit data analytics that have emerged as a result of these rapidly evolving different corporate systems and which have not been addressed. This paper highlights six general areas of such concerns and now provides a broad review and collection of additional critical ADA issues that challenge the public auditing profession today.

3.6.1 Research Questions

Many of the issues and sections reiterated similar research questions. Additional research questions are now presented that seem to be also important to answer for audit data analytics to succeed in gaining widespread practical acceptance. Also, quantification of many audit processes and judgements may be called for with the heightened use of advanced analytics and big data.

1. How can analytics methods be used to create accurate expectation models for generating predictions to compare with actual accounting numbers? How should allowable variances of predictions be chosen (Bumgarner and Vasarhelyi 2015)? Expectation models should be examined in greater depth

with the application of more advanced analytics. These more advanced approaches, combined with big data, may establish a narrower variance of prediction.

2. What properties make a particular ADA technique more or less appropriate for a particular audit function? There is a wide range of techniques appropriate for each audit phase, given the client particularities, environment, and industry. The categorization of appropriate techniques given certain client conditions is proposed as an External Audit Framework (EAA) in Appelbaum, Kogan, and Vasarhelyi 2016.
3. What types of "suspicion functions"[11] should be utilized in a preventive audit[12] as contrasted with transaction or account reviews? The weighting of characteristics of variables in linear suspicion functions may be impacted by ADAs such as expert systems, Bayesian Belief systems, probability models and Exceptional Exceptions (Issa, Brown-Liburd, and Kogan 2016).
4. How should the assurance function be reorganized to better use ADA? The assurance function is broader than that of financial statement auditing. Since assurance services should improve the quality of information for decision makers, the quality (relevance and reliability) of data is still

[11] A "suspicion function" is a linear multivariate equation that gives weights to characteristics of variables and analytical evidence to estimate its probability of being fallacious.

[12] Bumgarner and Vasarhelyi (2015) break down audit as retroactive and predictive. A predictive audit may be preventive (when a suspicion score is large, a transaction is blocked for review), or just predictive to set up a standard of comparison.

paramount. The assurance function may be reorganized in a broader format than the engagement, but standards must continue to be issued.

5. How should audit standards and processes be modified to enable and encourage the utilization of ADA? The standards should be modified to suggest techniques that are acceptable for each phase of the audit, given certain engagement contexts. For example, perhaps sampling should be modified for client engagements where 100% of the data is electronically collected and available to the auditor. In this context, ABE or Exceptional Exceptions (Issa, Brown-Liburd, and Kogan 2016) should be acceptable by the standard setters in lieu of sampling where appropriate. Additionally, the standards regarding data as audit evidence should also be examined in the context of electronic data and big data – external evidence may not be as reliable in this case (Appelbaum 2016; Brown-Liburd and Vasarhelyi 2015).
6. What is the proper way of validating expectation models for ADA? Should this validation be carried out for each audit client separately, or can it be extrapolated from one client to all the other clients in the same industry? Validation of models may be established over time by auditors for continuing clients and also for the auditors' own industry expertise. As part of interim activities, updated information could be fed into prescriptive analytical models that over time attain greater accuracy. The standards could also feasibly provide guidance specific for certain industries.

7. What additional verification processes would be desirable with the extant analytic technology? Verification processes and validation remain as open issues with ADA integration in the engagement. Over time, with continuing audit clients, it is likely that prescriptive analytics will become more precise.
8. How can the concept of "accuracy[13]" be defined for ADA? Is it necessary to encourage the use of substantive audit analytics? The concept of accuracy may be formally and quantitatively defined with the use of ADA. Auditor judgement is still necessary, even with advanced analytical techniques.

3.6.2 Evolution Towards Quantification of the Audit

Radical changes in analytics, information processing, and information distribution technologies have allowed assurance that can be continuous (Vasarhelyi and Halper 1991), predictive (Kuenkaikaew and Vasarhelyi 2013), prescriptive (Holsapple et al. 2014), and even facilitate automatic data correction (Kogan et al. 2014). These techniques are intrusive, create transparency, and maybe also some competitive impairment if all the details are disclosed, and generate substantive concerns by the auditee. The public good tradeoff of increased information disclosure versus economic interest of agents is a complex issue and its equilibrium may take many years to be reached, just to be disturbed by additional disruptive technologies.

[13] Acceptable relative error in engineering, materiality in accounting.

The increased amount of data available and the progressive ability to discover variances, understand aggregate content, and to predict trends has clearly created an equilibrium misbalance that is becoming larger and larger. Quantification can increase the value of information both internally and externally, but it decreases information asymmetry which is very threatening for agents (managers) and principals. A common thread of research questions relative to quantification were raised throughout this paper and are elaborated upon here:

- Do modern disclosure and statistical methodologies make it possible to, in certain cases, automate pre-set rules in order to perform procedures, derive results, and integrate these in a larger judgment? Such an approach is necessary for "close to the event continuous auditing" (Vasarhelyi and Halper, 1991) that is progressively been made necessary due to large electronic data streams exogenous and endogenous to the company.
- Research is needed on modern analytic methods, their applicability in different instances, their cumulative effect, their ability to be formalized, their classification (creation of taxonomies of analytic methods and data structures)[14], and their quantification. Traditional audit is backward looking due to the limitations of manual review and storage procedures. These modern analytic methods allow for the detection and prevention of propagation along downstream systems of potential faults (Kogan et al., 2014). These characteristics would force new corporate procedures of timely midstream

[14] The AICPA has created the Audit Data Standard (Zhang, Yang, & Appelbaum, 2015) to guide in the formalization of data to be received in the audit, its classification (into cycles), and its measurement.

error correction that do not exist in extant systems. These emerging procedures will be difficult to conceptualize from the point of view of "lines of defense" (IIA, 2013[15]; Freeman 2015; Chambers 2014), as they potentially make such lines blurred.

- If a midstream process detects faults and activates an error correction process that is a mix of human judgment and automatic correction, is this an audit or a control process? Does this distinction make sense in the modern world of automation?
- If a continuous audit layer detects "serious faults" (Vasarhelyi and Halper, 1991) and stops a system, is this layer a part of operations, control, or audit?

- Can audit findings and judgments be disclosed in more disaggregate manner with the use of drill-down technologies where the opinion would be rendered and broken down into sub-opinions and quantified in terms of probabilistic estimates (Chesley 1975, 1976, 1977)[16]. The issue of additional information disclosure in audit opinion is considered in the new PCAOB proposal and does not directly address the type of precision that disaggregation would allow. Turner (2014, p5) in the aforementioned comments to the PCAOB states "*it is clear that some oppose any disclosure of information not previously disclosed*

[15] "The tree lines of defense in effective risk management and control", White paper, The Institute of Internal auditors, January 2013.

[16] More detailed and quantitative audit reports are being progressively disclosed. For example, in the Netherlands (annual report of Aegon NV, 2015, p309) there is disclosure of the threshold of materiality EUR 65 million and the statement that "We agree with the audit committee that we would report to the misstatements identified during the audit about EUR 4 million (2014: EUR 4 million) as well as misstatements below that amount that, in our view, warranted reporting for qualitative reasons." Quantitative assessments are also made of coverage and other variables as well as a much more detailed discussion of governance controls and procedures.

by management. But such an approach defies common sense and is intended to obfuscate and avoid disclosing the information investors want. I urge the Board to reject such an approach as it will result in disclosures that are not worth the time or cost… investors wanted…information that is not "filtered through management" (adapted)." Improved stochastic estimates in disclosure, although not deterministic statements that create illusory comfort for the readers, may be the solution for this dilemma. Research here is urgently needed.

- Should quantitative guidelines be issued for ABE and its structures, and should within period results be disclosed as part of the auditor's report? A technological continuous audit allows for continuous monitoring and remarkable (not necessarily material) exception reporting. Should these exceptions be reported to all stakeholders (e.g. investors, suppliers, etc.) or only to select stakeholders? Should some of these exceptions be linked to smart contracts (Kosba et al. 2015) that automatically would execute a pre-agreed (e.g. covenant condition) action? A continuous assurance environment requires that events of substance, that can be predicted, be diagnosed and some action executed. As the combinatorics of these events is almost infinite, progressively more and more complex audit (and operational) judgments will be necessary, occupying auditors but necessarily changing their skill requirements (Tschakert et al. 2016).

3.7 Concluding Thoughts

This chapter contributes to the literature by discussing additional concerns facing the external audit profession as business moves towards big data and advanced analytics for many aspects of operations and decision making. These suggested research issues, along with various proposals towards greater use of big data and analytics will hopefully encourage and inspire ideas and research that is useful for professionals, regulators, and researchers. Although many concerns are reviewed, many are also not mentioned. It is expected that as research and findings evolve in this domain, some concerns will become less important while others many unexpectedly gain urgency. However, the emerging overall importance that big data and advanced analytics present to the public audit profession cannot be ignored.

In conclusion, big data and business analytics are dramatically changing the business environment and the capabilities of business processes. Business functions are changing, business capabilities are being added, anachronistic business functions are being eliminated, and most of all, processes are being substantially accelerated. The same has to happen to the external audit or assurance function, its rules need to be changed, its steps evolved, automation integrated into its basic processes, and its timing should become almost instantaneous in predictive, prescriptive and preventive analytical modes.

CHAPTER FOUR: SECURING BIG DATA PROVENANCE FOR AUDITORS: THE BIG DATA PROVENANCE BLACK BOX AS RELIABLE EVIDENCE

4.1 Big Data

Many client systems now are increasingly integrated with the cloud, the Internet of Things, and external data sources such as social media. Client data in the modern audit may exhibit large variety, high velocity, and enormous volume – big data (Cukier and Mayer-Schoenberger 2013). This data may originate from sensors, videos, audio files, tweets and other textual social media – all data types typically unfamiliar to an auditor (Warren et al. 2015). However, this big data provides almost limitless opportunities to the external auditor to utilize advanced analytics. According to extant analytics research (Holsapple, Lee-Post, and Pakath 2014; Lee et al. 2014; Delen and Demirkan 2012), big data should provide auditors the opportunity to conduct prescriptive analytics – that is, to apply techniques that computationally determine available actions and their consequences and/or alternatives, given the engagement's complexities, rules, and constraints (Lee et al. 2014).

Furthermore, this environment of Big Data (Vasarhelyi, Kogan, and Tuttle 2015), personal devices and the Internet of Things (IoT) (Atzori, Lera, and Morabito 2010; Domingos 2011; Dai and Vasarhelyi 2016) is progressively interconnecting with corporate systems.[1] The economics of hardware and software development are of very different nature than traditional systems. It is not inconceivable that analytic

[1] It is not surprising that this hybrid environment with numerous points of access and interconnections is a fertile ground for cyber-intrusion.

methods such as regression may be built into chips, including powerful explanatory software[2] that would provide interpretations of the results and recommend decisions for the user, in this case an auditor.

Advances in text interpretation, voice recognition, and video (picture) recognition would additionally expand the interconnected environment previously described. On another dimension, the latency of information and its processing systems are progressively reduced, mainly as the result of faster chips, interconnected devices, and automatic sensing of information. The traditional annual audit, or even quarterly report evaluation would have limited meaning in this world of real-time measurement. A progressive audit[3] by exception methodology would be required in this type of environment.

How can the availability of big data sets, both internally and externally to the enterprise, be utilized to enhance analytics? Can the extremely large amounts of data compensate for uncertain or, at times, lower quality of such data? There are some that argue that big data is meant to be messy (Cukier and Mayer-Schoenberger 2013). In cases where big data is of dubious origins or lacking audit trails, the standards currently would indicate that no amount could compensate for being poor, unreliable data – hence the eigtth research initiative which was also mentioned in Chapter Three:

[2] Byrnes (2015) has developed a clustering decision aid that can make decisions in the clustering interpretation process without human intervention. More sophisticated devices can be built into chips to accelerate and formalize this process and can benefit from standard interfaces and protocols.

[3] Montgomery (1913) already argued for a "continuous audit" that would provide progressive review results instead of the final audit opinion.

ISSUE 8: How can the provenance of external Big Data provide assurance as audit evidence?

In this big data environment with its many sources of information that would be novel for the audit profession to include in the examination, the standards regarding audit evidence may need to be discussed and possibly re-examined in the context of big data. Regardless of the source, the data should be reliable and verifiable. Table 14 outlines the challenges that big data poses to the current audit profession and suggests avenues of research:

Challenge of Big Data	**Recommendation**
How can the availability of big data sets be used to enhance analytics?	Research can suggest analytical techniques that take advantage of big data and evaluate how they improve audit effectiveness and/or efficiency.
Can the volume of data compensate for uncertain or lower quality of data?	Studies should be conducted that determine if there exists an upper threshold of data volume, exceeding which could compensate for lower data quality. A framework for data value should be generated.
How can the amount of audit evidence provided by analytics in a big data context be measured?	Research should re-examine the concept of whether evidence derived from analytics is “soft” and a quantitative reliability scoring system developed for all types of audit evidence. This score could then be integrated in the overall risk assessment.
How can big data evidence be aggregated with other types of audit evidence in a methodologically sound way?	This research question can be integrated with that of the data measurement system.
How can quantitative measures be used to provide support for the auditor’s judgment about the sufficiency of audit evidence?	This research question can be integrated with that of the data measurement system.
Alterability: How can the auditor be assured that the data has not been altered?	Research examining various tests for the assertion of accuracy in a big data context should be conducted.
Credibility: How can the auditor be assured of the controls surrounding the generation of big data external to the client?	Research examining/suggesting certain verifications of controls should be undertaken.
Completeness: How can the auditor verify that the big data is complete?	Research should be undertaken that can provide suggestions as to the verification of big data for the assertion of completeness.
Approvals: Should big data provide evidence of approvals/controls validations? Is this viable?	Studies of controls measurements of big data at all levels of generation and extraction

	should be conducted. For example process mining techniques (Jans et al, , 2014) can be used.
Ease of Use: Will big data require expertise to understand and extract and prepare for analysis?	What level of expertise should engagement staff attain to be competent in the modern audit engagement? This question is addressed later in this paper.
Clarity: Can this big data be replicated/re-performed/recalculated by the auditor?	Research should examine whether this is a viable test in a big data context and if so, how to perform it? This is the level of accuracy to be demanded from big data analytics. The concepts of materiality and relative error in the context of big data audit analytics should be examined in research

Table 14: Issues regarding big data as audit evidence, expanded from Brown-Liburd and Vasarhelyi, 2015

How can the amount of audit evidence provided by analytics in a big data context be measured? How can this evidence be aggregated with other types of audit evidence in a methodologically sound way? How can such quantitative measures be used to provide support for the auditor's judgement about the sufficiency of audit evidence? The entire standards of audit evidence may need to be reassessed and subsequently revised in this age of electronic and big data evidence (Appelbaum 2016; Brown-Liburd and Vasarhelyi 2015). Electronic and big data evidence often raise issues opposite of those assumed by the standards for paper-based documentation. As business processes now are very infrequently paper-driven, the standards on reliable evidence, which are derived from quality evidence of sufficient amount, may need to be revised to provide a more quantitative measure of quality vs. quantity in an IT audit.

4.2 Discussion of Big Data and Auditing

Big Data has become the new business currency (CompTIA 2015). To this end, businesses are now collecting more data than they have in the past 2000 years (Warren, Moffitt and Byrnes 2015). These businesses regard big data as a potential firm asset (Warren et al. 2015; Brown, Chui, and Manyika 2011) and have been reported to have attained five to six percent gains in productivity from analysis of this data (Brynjolfsson, Hammerbacher, and Stevens 2011). There is an enormous quantity of data now available in many forms from many different sources that is being generated very quickly - 2.5 quintillion bytes of data are being generated daily (IBM 2015; Jagadish et al 2014) - a Big Data deluge (Hey and Trefethen 2003). Most of these datasets are unstructured, derived from social media, sensors and the Internet of Things (IoT)(Bauer and Schreckling 2013). As such, Big Data is dynamic data with volume, variety, and velocity (Laney 2001), and more recently veracity (IBM 2012). Big Data may be defined as the large flows of widely differing data and the aggregation of datasets that cannot be processed using traditional database management tools (Polato, Goldman, and Kon 2014; Mittal 2013; Zikopoulos and Eaton 2011). Furthermore, the origin and treatments of these datasets are largely unknown as they often originate outside of the business that is absorbing and analyzing it (Taylor, Haggerty, Gresty and Hegarty 2010; Tan 2007; Cui and Widom 2003).

For decision makers, researchers, auditors, and regulators, the ability to verify the accuracy of information is of paramount importance (Liao and Squicciarini 2015; Ikeda, Park, and Widom 2011; Li, Roge, Rydl, and Hughes 2007; Nearon 2005; Alles, Kogan, and Vasarhelyi 2002; Elliott 2002; Elliott 1997). External auditors may be interested in Big Data for two reasons: one, their clients may be utilizing Big Data for decision making and accounting judgements that could materially affect the financial statements if the data is flawed; and secondly, auditors themselves may want to access Big Data sources for industry and client assessment, risk analysis, confirmations, and reasonableness tests – if the data is reliable.

The audit standards (Public Company Accounting Oversight Board [PCAOB] 2010, AS No. 15, American Institute of Certified Public Accountants [AICPA] 2012, SAS 122; International Auditing and Assurance Standards Board [IAASB] 2009, ISA 500) specify that external sources of evidence and information are generally more reliable for verification. However, Big Data potentially poses an opposite situation: due to its possible lack of provenance and veracity, it could be a less reliable source of evidence for auditors. Big Data may not be trustworthy if the organization utilizing it has not employed certain procedures to address its risks (Zhang, Yang, and Appelbaum 2015; Mittal 2013). Basically with Big Data, much of the innovation has been directed towards processing and analyzing this data of such volume, variety, and velocity and not tracing its veracity, or origins and transformations (Liao and Squicciarini 2015). Until very recently, little attention has been paid to the Provenance[4] of this Big Data, its pedigree or lineage (Liao and

[4] Provenance traditionally has meant the chronology of the ownership, custody or location of a historical object. The term was originally mostly used in art, but is now used in a number of domains such as archaeology, paleontology, archives, manuscripts, printed books, medical sciences, and computing. "The primary purpose of tracing the provenance of an object or entity is normally to provide contextual and circumstantial evidence for its original production or discovery, by establishing, as far as practicable, its later history, especially the sequences of its formal ownership, custody, and places of storage. The practice has a particular value in helping authenticate objects. Comparative techniques, expert opinions, and the results of scientific tests may also be

Squicciarini 2015: Ikeda et al 2011).

Big Data, due to its volume and velocity, has compelled business organizations to utilize the cloud for data storage and enterprise applications (Polato et al 2014). Big Data, due to its immense volume, great variety of format, and streaming velocity of occurrence has forced numerous firms to utilize applications such as Hadoop MapReduce to process and prepare the data in a form that is manageable for analysis and understanding (Akoush, Sohan and Hopper 2013; Lin and Ryaboy 2013; Dean and Ghemawat 2008). However, both the cloud and MapReduce processing create additional challenges to the auditor for evidence verification (Cohen and Acharya 2014; Polato et al 2014; Lin and Ryaboy 2013). The Cloud is a data repository that resides outside of the business enterprise or cloud client, the result of which is that the enterprise has partially lost control of the data in an environment where provenance tracking is challenging. Hadoop and MapReduce process the streams of data and may alter and transform it without complete tracking of these alterations. For an enterprise processing Big Data with a Hadoop platform in the Cloud, these provenance issues could be magnified. Audit techniques should take into account the impact of this reliance on messy Big Data by the client. This Big Data may not be providing verifiable evidence for auditors and regulators, particularly if this data materially impacts the financial statements.

The auditor, whether internal or external, should be able to access the desired

used to these ends, but establishing provenance is essentially a matter of documentation". Extracted from https://en.wikipedia.org/wiki/Provenance level of provenance of the electronic information under examination, and this provenance tracking should be secure and trustworthy (Bates, Mood, Valafar, and Butler 2013; McDaniel et al 2010; Hasan, Sion, and Winslett 2009; Braun, Shinnar, and Seltzer 2008). The internal auditor could be utilizing big data from sensor streams and social media texts to perform efficiency and fraud auditing more efficiently and effectively (Warren et al 2015). As such, the origins and paths of lifecycle of this data should be verifiable by the auditor and this recording of its lifecycle, the data provenance, should be secure and unalterable. Similarly, external auditors could access Big Data in many forms, primarily from social media and the web, for example to augment the initial client evaluation decision, to verify the client's fair value assessment of intangible assets, or to evaluate the determination of going concern (Warren et al 2015).

To summarize, it is envisioned that the external auditor would directly access Big Data to enhance the following typical audit phases:

1. To supplement the auditor's industry and client knowledge acquisition during the Engagement Phase
2. To supplement the auditor in the risk assessment process of the Audit Planning Phase, similar to the Engagement Phase.
3. As part of Substantive Testing, particularly if re-performing client calculations and analyses that utilized information derived from Big Data. For example, verifying the client's Fair Value assessment of intangible assets that has been partially based on social media information is one task that would require the

auditor to access Big Data.

4. During the review stage, the auditor may want to view all the audit results in a greater context and in a comparative sense against the client's own industry and associated internet media. Critical to this analysis would be any direct social media or macro-economic/demographic Big Data that would indicate a probable Going Concern issue.
5. Big Data may also enhance the auditor's knowledge regarding the client in the Continuous Activities phase, similar to the Engagement and Planning phases.

Big Data could expand the auditor's client and industry knowledge beyond that provided from the client's own data. Evidence collection in this Big Data scenario could not only assist in traditional financial statement verification but also enhance auditor knowledge for client assessment.

Essentially, the traditional view of audit evidence collection may no longer be sufficient in this more advanced technical business environment (Brown-Liburd and Vasarhelyi 2015). The customary characteristics that define traditional audit evidence may not be adequate, and has been proposed as a future research issue (Brown and Vasarhelyi 2015). Previously, when the bulk of electronic data was internally generated and quantitative, provenance information was readily available to auditors via system log files (Caster and Verardo 2007; Cerullo and Cerullo 2003). In contrast, Big Data may not be internally generated and most likely has been processed outside the client. The provenance tracking that is missing for many Big Data and cloud systems would appear to challenge the long-held view in the audit profession that external data equals reliable data.

To expand upon this concern, the purpose of this chapter is to discuss the challenge of provenance evidence verification facing the auditor in the current electronic Big Data business environment, to identify the current gaps in the audit and systems research regarding secure Big Data provenance, and to propose a model and direction for future research – the Big Data Provenance Black Box. This Introduction is followed by a review of the Auditing Standards on Evidence Collection, where the evidence attributes are discussed and issues of digital evidence collection, with an emphasis towards external auditors, are highlighted. These attributes and evidence collection issues will shape the remainder of this discussion. The third section offers an overview of Provenance collection, emphasizing security. The fourth section discusses Hadoop/MapReduce and Hadoop in the Cloud and their impact on reliable evidence collection. The Big Data Provenance Black Box is proposed next, and the final section offers a conclusion and commentary on areas for future research regarding evidence collection in the current Big Data business environment and the external auditor.

4.3 The Auditing Standards on Evidence Collection

The main purpose of the work conducted by an auditor in an external engagement is to obtain reasonable assurance that the client's financial statements are basically free from material misstatements and to subsequently express an opinion regarding these financial statements in the auditor's report. To accomplish this task, the auditor must design and perform audit procedures to obtain sufficient appropriate evidence; furthermore, the Audit Standards require auditors to examine physical evidence as part of the risk assessment process (PCAOB 2010, AS 15; AICPA 2012, SAS 122; IAASB 2009, ISA 500).

Additionally, the Sarbanes-Oxley Act (SOX) demands that public auditing firms maintain the provenance of an audit report (and all of its supporting information) for at least seven years after its issuance (United States Public Law No. 107-204; Tsai et al 2007). The Sarbanes-Oxley Act also mandates that auditors verify the accuracy of the information or evidence that forms the basis of their audit opinion. Management also needs to be able to audit and verify each step of every transaction, with all its data inflows and outflows. The client's document management, access to data, and storage of information must provide auditing (vouching, verifying, and tracing) capabilities (Li et al 2007). As such, many public companies have sought to reduce compliance costs by collecting data in a real-time fashion to provide continuous monitoring of 100% of the transactions.

Audit evidence is all the information used by the auditors to form the audit opinion (PCAOB, 2010, AS 15). This audit evidence must be both sufficient and appropriate, the degree of each is determined by the other (see Figure 1). Sufficiency is the measure of the quantity, the amount of which is determined by Detection Risk determined by the auditor and the level of quality of the evidence, or it's Appropriateness (PCAOB 2010, AS 15). Appropriateness is the measure of Relevance (what does the evidence tell the auditor) and Reliability (can the auditor trust the evidence)? Basically, if the underlying information is not reliable and its provenance or lineage isn't verifiable, then more evidence will need to be collected and reviewed – to a certain degree. Poor quality evidence cannot always be compensated by collecting a larger amount (PCAOB 2010, AS 15). If the evidence is relevant and reliable, possessing trustworthy provenance, then the auditor can proceed confidently with substantive testing and other analytical procedures (PCAOB 2010, AS 15). Traditionally, much of this evidence has been paper, observations, inquiries, and other physical formats. As shown in Figure 25 the aspect of Appropriateness is quite significant to the determination of Detection Risk.

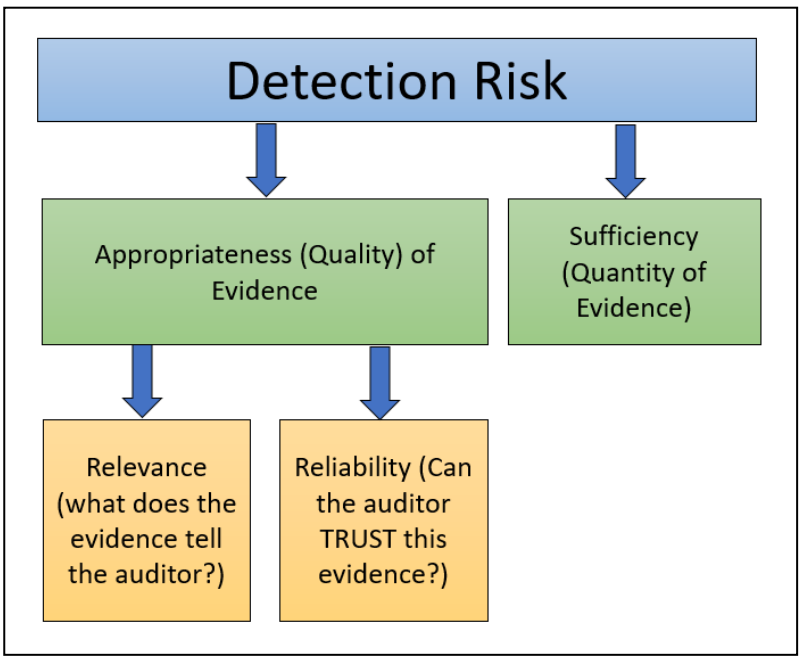

Figure 25: Depiction of the role of Appropriateness and Reliability of evidence in Detection Risk

However, in today's complex IT environment and Big Data, the nature and competence of this audit evidence has changed (Brown and Vasarhelyi 2015; Caster and Verardo 2007; Nearon 2005). Every phase of a transaction is computer generated and recorded and can only be verified electronically. For example, every phase of a purchase or a sale may occur within the electronic system. Or, with additional information available from external Big Data, intangible assets might be partially valued by the client from information derived from text analysis of aggregated tweets and web scraping of social media. With more than 90% of these records in easily alterable digital formats that possess many iterations and possibilities, provenance of data sources and provenance of log files become of paramount importance (Nearon 2005). To summarize and expand, Table 15 displays the following differences that exist between paper evidence and electronic evidence (Brown-Liburd and Vasarhelyi 2015; Colbert and Smalling 2011; Ratcliffe and Munter 2002):

Evidence Characteristics:	**Paper Evidence:**	**Electronic Evidence:**
Alterability: easily altered evidence lacks credibility; evidence should be difficult to alter	Difficult to alter without detection	Alterations may be difficult to detect without performing specifically designed tests
Prima facie credibility: SAS 80 establishes a hierarchy of credibility –outside sources enhance credibility when independent of the client and confirmable	Outside sources of paper and documentary evidence and submitted directly to the auditor enhance credibility; inside sources of paper evidence that have been reviewed and processed by outsiders is also reliable	An electronic document derives its credibility primarily from the controls within the system. Outside electronic documentation/data is missing the assurance of system controls that the document or data is not fraudulent or altered
Completeness of documents: All essential terms of a transaction are verifiable	Typically all essential terms are included on its surface in a text/human readable form	An electronic system may substitute codes or cross-references to other data files that may not be accessible
Evidence of approvals: This essential aspect of internal controls should be easily verifiable and transparent	Approvals integrated into paper documentation add to completeness	Electronic approvals may be similarly integrated, but need additional verification

Ease of use: Simplicity of application and access encourages compliance	Paper evidence can usually be evaluated without the use of additional tools and/or skills	Electronic evidence may require extraction of data by an expert
Clarity: competent evidence should allow for the same re-performance and conclusions by other auditors	The nature of paper documentation is readily clear	The nature of electronic evidence is not always so clear, particularly in the absence of appropriate controls

Table 15: Review of Evidence Characteristics, adapted from Brown and Vasarhelyi 2015; Colbert and Smalling 2011; and Ratcliffe and Munter 2002

The implications for electronic accounting data and evidence collection are substantially different from that of manual, paper-based examination. Many of the characteristics that are strengths with paper-based evidence pose issues for electronic evidence. It could be said that technology has weakened a number of traditional forms of audit evidence (Caster and Verardo 2007). Whereas paper documentation is considered not to be easily altered, electronic data may be easily changed and these alterations might not be detected, absent the appropriate controls. In paper-based evidence collection, sources that are verified external to the client are considered to be highly reliable, whereas external electronic evidence is difficult to verify for veracity, origin, and reliability. External data also frequently lacks evidence of approvals and signatures. Paper-based evidence is easy to evaluate and understand, whereas electronic data and evidence may require a high level of technical expertise of the auditor. Also, whereas manual paper-based information is competent for re-performance and re-calculation, electronic evidence may require additional complex procedures due to its random and dynamic condition. These five characteristics will assist in the evaluation of Big Data and the suggested provenance collection system in the remainder of the paper.

Statement of Auditing Standards No. 80 (SAS 80) *Amendment to Statement on Auditing Standards No. 31, Evidential Matter* was released to provide guidance regarding audit evidence collection in electronic environments (ASB 1996). SAS 80 clarifies that tests of IT controls, together with substantive testing, may provide sufficient evidence to form an audit opinion if the client's reliance on IT is so great that detection risk cannot be limited to substantive testing alone (Auditing Standards Board [ASB] 1996, SAS No. 80). IT controls may be examined by inspection of log file activity for compliance verification. Log files record the dynamics, the activity flows and events in a system. A log file will record the data or transaction origin if this information was provided and any subsequent changes with time/location/authorization/actor stamps and identifiers (Accorsi 2006). Logging in fact has typically been recognized as the recording of significant events that may need to be identified in a future audit. These log entries should be considered as evidence of origins, authorizations, permutations, alterations, IP addresses, and time strings (Vaughan, Jia, Mazurak, and Zdancewic 2008). Log files are also considered to be the starting point for process mining (Jans, Alles, and Vasarhelyi 2010; van der Aalst, van Hee, van Werf, and Verdonk 2010), where the systemic, reliable and trustworthy recording of events and data (business provenance) is required. Additional future research and discussion could focus on how provenance log files may provide sufficient evidence for internal control compliance evaluations in an electronic or Big Data environment.

Nearon (2005) proposed that an appropriately skeptical auditor should inquire as follows regarding electronic evidence, log files, and IT controls:

- Is the electronic evidence subject to alteration without an audit trail or evidence of this change?
- Is there an audit trail that clearly ties the digital evidence back to the initiating entry or data origin? Or, can this trail lead forward to the point of inclusion on the face of the financial statements?
- Does the electronic evidence include metadata that identifies who made the entry and when?
- What are the controls designed to prevent unauthorized changes to the digital evidence after it was created?
- Who has or had access rights to change the digital evidence?
- How does the auditor know that the digital evidence hasn't been intentionally altered?
- Has the audit logging process been configured to record all access attempts, whether successful or not?
- Have the audit logs been reviewed independently?
- Has the continuity of logs been maintained and any gaps justified?
- Have the logs been frequently copied to off-line, read only media and stored in a separate secure location, inaccessible to those who might be motivated to change it?
- Has the access to the logs and their security settings been recorded, and limited to only authorized persons?

All of these questions could potentially be satisfied with an appropriate secure provenance system, which will be discussed in the sections that follow. Basically, business provenance provides assurance of traceability, verifying the lineage of the event or transaction. With the assurance provided by the reliable and trustworthy recording of event logs and audit trails, known as provenance of process flows, the auditor can embark on a risk assessment analysis based on a secure foundation of accurate accounting data, event log files and process flows.

4.4 Data Provenance and Evidence

4.4.1 Data Provenance

Provenance by definition means origin and lineage, and is used quite extensively in the arts, antiques, and scientific domains to describe lineage or ownership of different items (Moreau et al 2008a, 2008b, 2008c). When applied to data, provenance may be metadata or log files/audit trails pertaining to the lineage of a data event, capturing and recording its origins, derivations, and transformations and has been used extensively in the sciences (Bose and Frew 2005; Moreau et al 2008c; Simmhan, Plale and Gannon 2005a, 2005b). As businesses increasingly depend on data from sources outside the firm, such as Big Data, the need for provenance of this data grows exponentially (Cheah and Plale 2012).

As the available data has become larger, i.e. Big Data, the analysis required to achieve knowledge discovery requires more complex and distributed processing (Crawl, Wang, and Altintas 2011a; Davidson and Freire 2008; Frew, Metzger and Slaughter 2008). Therefore, it is quite possible that the originating data could have been entirely different from the data that the organization now possesses, due to preprocessing applications (Cheney, Chiticariu, and Tan 2009; Glavic and Dittrech 2007; Scheidegger et al. 2008; Simmhan, Plale, and Gannon 2005a). Hence, provenance is essential to the business domain as it may be used to provide an audit trail for regulatory and audit engagement purposes (Simmhan et al 2005a, 2005b). For the purposes of this paper, data provenance is considered to be all the information that assists in determining the origin, derivations, and transformations of a data product or dataset (files, tables, process flows, log files, virtual collections) (Cheah and Plale 2012). Two main features of data provenance are the originating data product itself and the process flows that record the activity and locate points of transformations of the originating data product to its current form (Ikeda and Widom 2010; Tsai et al 2007).

Data provenance can be available explicitly or deduced indirectly. The explicit model, or data-based model, collects lineage metadata about the data and transformations directly. A provenance Directed Acyclic Graph (DAG) [6] is directly associated with the data product whose lineage it describes. The indirect model, or process oriented model, describes the deriving processes that contribute to a dataset's existence.

Provenance may also be fine grained (explicit and detailed) or course grained (deduced and processed through a workflow) (Tsai et al 2007). The size of provenance

---

[6] A Directed Acyclic Graph (DAG) is a design from computer science that models a wide variety of activities or process flows. The DAG consists of the following elements: Nodes, which represent objects or points of data; Directed Edges which are directional arrows or edges from one node to another; A Root Node, which has no parents and only children; and Leaf Nodes which have no children. Arrows in a DAG may not form a cycle, where these arrows illustrate the basis. A DAG may be considered to be a tree like data structure, similar to decision trees. – extracted from http://ericsink.com/vcbe/html/directed_acyclic_graphs.html

information may exceed that of the dataset's and the storage costs may be substantial. The storage location and format of the provenance should also be determined by the frequency and application of use. The granularity (and hence the cost) of the provenance to be recorded will depend on the inherent risk of the business cycle, the origins of the data (internal/external), the type of dataset (structured/unstructured) and the impact or potential materiality of the dataset on the financial statements. Table 16 summarizes the provenance types generally applicable to the audit examination tasks:

Purpose/Audit Task:	**Provenance Type:**	**Qualities:**
Internal Controls Verification/ re-performance	Coarse-grained	Work flows or process flows based; data at schema level; DAG models
Evidence Collection/Verification; recalculations	Fine-grained	Data elements/metadata; DAG models

Table 16: Review of generally suggested provenance types per audit task

The business domain has typically worked with organized, quantitative and mostly internally generated data, where the structure and semantics of the data is organization-wide. However, many businesses now collecting and analyzing data that are messy and unstructured, whose issues are further compounded by its aggregation to a data warehouse (CompTIA 2015). Basically, the data is required to be extracted, cleansed, and transformed from many different operational databases and external sources before it is placed in a data warehouse or a cloud. Provenance is also essential in a warehouse environment, as warehouse data is built upon layers of data views, with one layer derived from layers below it, and where lineage information is essential for vouching and tracing. This warehouse provenance data product and its transformations may be conceptualized graphically as a DAG with nodes representing the different iterations of the data product and with the edges revealing each of the transformation processes.

Goble (2002) summarized the feasible applications for provenance information and that research has been adopted and modified in this paper to the external audit domain as follows:

- Data Quality: Lineage can estimate and verify data **quality** and data **reliability** based on the source information and transformations (Simmhan et al 2005a).. The level of data included in the provenance determines the extent to which the quality can be estimated – the more fine grained (detailed) the provenance, the more precise the estimation of data quality. The more coarse the provenance (summary level), the less detailed the estimation. The granularity of provenance to be recorded may vary based on the inherent risk of the business cycle, the origins of the data (internal/external), the type of dataset (structured/unstructured) and the impact or potential materiality of the dataset on the financial statements.
- Audit Trail: Provenance can provide a means by which to audit the **veracity** of the data and the process by which it evolved. This information is important for accounting and auditing purposes, particularly for data that is ambiguous. The standards stipulate that uncertain evidence or data must be thoroughly examined with substantive procedures such as re-performance, recalculation, trend analysis, analytical procedures, and vouching/tracing (ASB 1996, AS

80). Lineage can help identify any exceptions that took place in data creation. Provenance can also be used to back track and identify the source of errors and violations of controls (Galhardas, Florescu, Shasha, Simon, and Saita 2001).

- Replication Recipes: Detailed or fine-grained provenance can allow repetition of data derivation and be a recipe for its re-performance or recalculation. Re-performance and recalculation are integral procedures for most audits of financial statements. With provenance, the auditor can vouch and trace from the dataset origin to the face of the financial statement and vice versa. Many current applications of provenance have adopted XML for representing lineage information (Bose & Frew, 2004). As a suggestion for future research, XBRL, as an XML derivative, may present possibilities to the business domain as a provenance metadata standard, particularly since public companies currently are required to prepare their financial statements in XBRL.
- Attribution: Pedigree or lineage can help determine or verify ownership of the source data used to generate certain estimates or calculations. An auditor can verify the creators of intellectual property and copyrights or look at the lineage chain to see who has had access. Lineage is also the means by which citations are tracked in the academic publications domain (Cameron 2003). Provenance can also be used to assign liability in case of errors in the dataset (Cameron 2003).
- Informational: A more generic use of provenance is as a metadata categorization that may be utilized for queries, with the trail of any particular query available for re-performance, avoiding duplication of effort. Annotations

that accompany the provenance may help interpret the data in the context required, particularly for archived data that is accessed long after it was generated (Simmhan et al 2005a, 2005b).

Actually, without assurance that this data provenance has been collected and maintained securely, the audit records of the origins and transformations of this data is suspect (Cheah and Plale 2012; van der Aalst et al 2010; Buneman, Khanna and Tan 2007, 2001, 2000). The use of any provenance as a basis for decision making, whether by the client or auditor, depends on the trustworthiness of that provenance information itself (Bier 2013; Aldeco-Perez and Moreau 2010; Simmhan et al 2005a, 2005b). There should be assurances that the provenance information was not tampered with and securing provenance with digital signatures has been a common solution (Aldeco-Perez and Moreau 2010; Simmhan et al 2005a, 2005b). Securing provenance information will significantly enhance its usefulness and value for auditors as a reliable source of examination evidence and accounting data.

4.4.2 Secure Data Provenance

Provenance has been recognized, due to its ability to track causal dependencies between data and events that explain the data's current state, as a means to achieve information accountability (Aldeco-Perez and Moreau 2010; Moreau et al 2008b; Weitzner et al 2008). Provenance provides transparency of the datasets it reflects and is auditable, allowing auditors to decide whether information is credible or has been used in the proper way. However, the integrity of this provenance information and its graphs are critical to guaranteeing the quality of a data provenance based audit.

Basically, the auditor should be able to verify that the information tracking the subject datasets has not been altered itself. Most research to date has suggested digital signatures to be the most feasible means of securing the provenance documentation (Bier 2013; Aldeco-Perez and Moreau 2010; Accorsi 2009; Accorsi 2006; Simmhan et al 2005a, 2005b). The provenance information flows should be recorded securely in these four stages in order to guarantee a correct audit report (Aldeco-Perez and Moreau 2010):

- Recording of any process documentations in which influential components make assertions about the actions they perform on the dataset, in addition to the alterations
- Storage of the provenance information in which it is continually stored in a Secure Provenance Repository separately located with highly enforced access controls and is read-only
- Querying of the provenance information should also be recorded
- Analysis of provenance information should be recorded, which provides the basis for the audit report

If the provenance data and DAGs are secured via digital signatures at the formation, recording, storage, querying, and analysis stages, the provenance data may be regarded as reliable for auditors (Accorsi 2009; Alles et al 2004). With the use of digital signatures, security is assured in the transmission and storage phases. In the transmission phase, origin authentication, message confidentiality, message integrity, message uniqueness, and reliable delivery are assured with digital signatures.

Similarly, in the storage phase, entry accountability, entry integrity, entry confidentiality, and tamper prevention are assured.

With digital signatures, a small change to the original data results in a huge difference to the hashed message (digital signature). It is computationally impossible to create two different documents that have the same digest; so if one document is altered, it would be impossible to create another document with the exact same digital signature. A digital signature does not reveal any information about the content of the provenance data itself, only if the content has been altered (Alles et al 2004). With digital signatures, not only is the transmission and storage of provenance records secure, but this security itself is assured. With digital signatures, the provenance information cannot be thwarted.

The ability of secure provenance to satisfy the requirements of audit evidence that were discussed in Section 2 from the Audit Standards are shown as follows (Table 17):

Evidence Characteristics:	Paper Evidence:	Electronic Evidence:	Secure Data Provenance:
Difficult to alter	√		√
Credible	√	√ for internal data	√
Complete	√		√
Evidence of approvals	√		√
Easy to use	√		
Clear	√		√

Table 17: Summary of satisfaction of audit evidence characteristics by evidence type

The summary table in Figure Four summarizes the information from Figure 2, extended with the attributes of a secure data provenance storage system using digital signatures. Although secure data provenance comes close to meeting the attributes of audit evidence as required by the audit standards, it is not considered to rank highly for ease of use generally. For auditors to navigate a secure provenance data warehouse, applications would need to have been scripted that would be interactive and provide a simple interface. Such applications have been proposed by academics using Python, Perl, or Matlab (Simmhan et al 2005a, 2005b).

However, to date there has not been research published specifically about secure provenance of Big Data in Hadoop. This may be due to the rapidly expanding exposure and availability of big data, in which common applications such as MapReduce and high capacity storage locations such as the Cloud have neglected provenance issues until recently (Polato et al 2014). There are many studies of Hadoop or MapReduce in the area of Big Data, but only a few that discuss data provenance in Big Data or Hadoop (Chen and Plale 2015; Imran, Agrawal, Walker, and Gomes 2014; Akoush et al 2013; Che, Safran, and Peng 2013; Goshal and Plale 2013; Crawl, Wang, and Altintas 2011; Park, Ikeda and Widom 2011; Simmhan et al 2005b). Furthermore, none of the studies provide for a secure form of data provenance in Big Data applications (Ikeda, Park, and Widom 2011; Margo and Smogor 2010; Aggarwal 2009; Bao, Cohen-Boulakia, Davidson, Eyal, and Khanna 2009; Muniswamy-Reddy et al 2009; Souiah, Francalanza, and Sassone 2009; Cohen-Boulakia, Biton, Cohen, and Davidson 2008; Freire, Koop, Santos and Silva 2008; Buneman and Tan 2007; Davidson et al 2007; Glavich and Dittrich 2007; Muniswamy-Reddy, Holland, Braun, and Seltzer 2006; Simmhan et al 2005a, 2005b; Tan 2004; Buneman, Khanna and Tan 2001). Basically, if the provenance information about the Big Data cannot be stored securely, there is no point in collecting it for auditing purposes. Without security measures, the data provenance recording is not reliable (Buneman and Davidson 2010). For auditors, unreliable information equals poor quality evidence.

4.5 Hadoop/MapReduce and the Cloud

4.5.1 Hadoop/MapReduce

In the realm of Big Data, MapReduce applications such as open source Hadoop have been widely adopted (Akoush, Sohan and Hopper 2013; Dean and Ghemawat 2008). Hadoop as a MapReduce agent has become synonymous with Big Data processing and analysis (Crawl et al 2011), particularly in larger public companies (CompTIA 2015). Hadoop was designed as an open source software framework that would provide a scalable distributed storage and parallel processing system for structured and unstructured Big Data sets (Cohen and Acharya 2014). If an internal or external auditor is working with Big Data, most likely he/she will be referring to datasets that have been processed with Hadoop. Many social media sources and aggregators of Big Data, such as Facebook, Twitter, Yahoo, and Google employ various forms of Hadoop or MapReduce (Lin and Ryaboy 2013; Patil 2012; Hammerbacher 2009).

Not only do these social media data generators utilize Hadoop and MapReduce, much of their qualitative, textual, video and audio feeds must be transformed and integrated before analysis (Lin and Ryaboy 2013). These processes may have altered the data and may not have been completely recorded or logged, unless provenance collecting applications were added. Furthermore, Twitter, which has become a predominant social media source for business promotion, customer service, political campaigning, medical services, health care, marketing, and stock market prediction (Chu, Gianvecchio, Wang, and Jajodia 2012; Bollen, Mao, and Zeng 2011; Hughes and Palen 2009), is plagued with issues of fraudulent accounts and spam campaigns whose origins are not clear/traceable (Cresci, Di Pietro, Petrocchi, Spognardi, and Tesconi 2015: Duncan 2015; Chu et al 2012; Castillo, Mendoza, and Poblete 2011; Thomas, Grier, Song, and Paxson 2011). Why is this important for auditors? Depending on the client industry and business cycle, Twitter data and other social media sources may have been used by the client in its analytics to gain additional insights beyond mere quantitative analysis (Lin and Ryaboy 2013; Bollen et al 2011). If the results of these analytics contribute to information that is material to the financial statements, then auditors should be concerned about the provenance of the contributing social media Big Data, as the risk of material misstatement has increased.

Hadoop consists of two functions: Map and Reduce. The user-provided Map function reads, filters, and transforms data from an input file, creating a set of intermediate records. These intermediate records are then usually split via a certain hash function into different buckets. Then the user provided Reduce function processes and combines all of the intermediate records associated with that hash value into new records which are written into parallel output files. Essentially the system splits large data sets into smaller pieces, distributes them to as many output files as possible, and then processes the data in each parallel folder so that it is tightly aggregated (Cohen and Acharya 2014). Processing speed and data replication were the core goals behind Hadoop's evolution, with provenance and security a secondary concern. Programs developed with the Hadoop model are parallel because there are no inter-key data dependencies. As such, MapReduce is tolerant of system failures as problematic functions can be restarted independently of the other parallel operations. MapReduce functions are usually expressed as a series of jobs creating a computational workflow. Provenance metadata are captured only at two main points within the core Hadoop platform, unless there have been additional specific provenance process applications added to the Hadoop software (Cohen and Acharya 2014).

Provenance metadata in the basic Hadoop are captured at the storage level and at the resource management level (Alabi, Beckman, Dark, and Springer 2015). The storage level metadata captures such information as file location, ownership settings, file type, permissions settings, and transaction history – all useful information for provenance. The resource management collects and tracks the data provenance related to the application of Hadoop, but at two points only (Alabi et al 2015). Therefore, much of the current research in Hadoop provenance is related to enhancing another aspect of provenance, the tracking and the lineage of the Hadoop application workflows (Alabi et al 2015; Akoush et al 2013). This additional course grained provenance serves the purpose for tracing and vouching the data outputs back to its associated input activities and origins, and vice versa, for the detection of data alterations or any type of suspicious activity.

As can be imagined, thecomplex Map/Reduce processes could result in an even more extensive provenance, larger than the workflow that it records and resulting in significant overhead; therefore current research has been focused on establishing feasible provenance collection in Hadoop (Alabi et al 2015). For example, one extension of Hadoop that was developed to support provenance capture and tracing for workflows of MapReduce jobs is Reduce and Map Provenance or RAMP (Park et al. 2011; Ikeda et al 2011). However, there was a fairly large runtime overhead of 76% on unstructured Twitter data. Another study presented an application of MapReduce in Kepler[7], a Kepler+Hadoop framework, to record provenance of workflows (Crawl et al. 2011). However, word count tests took 2.5x longer to execute when the provenance capture was enabled (Crawl et al 2011).

A more recent application of provenance in Hadoop is HadoopProv (Akoush et al 2013). HadoopProv was designed as a modification of Hadoop that takes advantage of the metadata that Hadoop captures while also tracking lineage of data at the process log level. The authors claim that provenance capture overheads are reduced by treating the Map and Reduce phases separately and deferring construction of the provenance Directed Acyclic Graph (DAG) to the query stage. HadoopProv was also designed to capture provenance at the record level, and this level of fine grained tracking allows for incremental process and log analysis. The temporal overhead of HadoopProv was 10% on a typical MapReduce workload (Akoush et al 2013). In all three approaches,

[7] Kepler is an open source software application for the modeling and processes of scientific data, see https://code.kepler-project.org/code/kepler-docs/trunk/outreach/documentation/shipping/2.5/getting-started-guide.pdf.

security measures of the provenance files were suggested by the authors as an area for future research.

4.5.2 Hadoop/MapReduce in the Cloud

Further compounding the issue of feasible provenance collection of Big Data is the recent migration of Hadoop platforms to the Cloud[8] (Olavsrud 2016). The Cloud has become a popular pay-as-you-go location for data storage, due to its flexibility and scalability (Assuncao, Calheiros, Bianchi, Netto, and Buyya 2014). Clouds are known for their ability to scale dynamically upward or downward depending on demand and workload. Hadoop and other Map/Reduce systems have also been established with Cloud providers as Platform as a Service (PaaS). However, the Cloud is perceived as being insecure (O'Driscoll, Daugalaite, and Sleator 2013; Armbrust et al 2010), providing scanty locational provenance as a result of this scalability and flexibility. Clouds are generally untrusted since the guarantees provided regarding data transformations and locations are minimal (Sakka, Defude, and Tellez 2010). Furthermore, most cloud providers offer clients little capability on data, application, and service interoperability. Most cloud storage services are not designed to effectively and efficiently store provenance data, due to the cyclic nature of provenance - its need to be stored separately yet linked to the data objects

---

[8] According to the National Institute of Standards (NIST), "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models". The five essential characteristics are: On-demand self-service, Broad network access, Resource pooling, Rapid elasticity, and Measured service. The three service models are Software as Service(SaaS), Platform as Service(PaaS), and Infrastructure as a Service (IaaS). The four deployment models are as Private Cloud, Community Cloud, Public Cloud, and Hybrid Cloud. – extracted from http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

(Muniswamy-Reddy and Seltzer, 2006). Currently, provenance of the Cloud persists as an open research problem (Assuncau et al 2014). For auditors, the use of the Cloud for either processing or storage of Big Data by a client may likely increase the risk that the relevant data is not reliable as audit evidence, due to the minimal provenance of transactions.

As discussed in Section Two, an appropriately skeptical auditor should inquire as follows regarding Big Data electronic evidence, log files, and IT controls in the core Hadoop platform or Hadoop in the Cloud Big Data context:

- Is the Big Data electronic evidence subject to alteration without an audit trail or evidence of this change? – Quite possibly the data has been altered in core Hadoop with minimal provenance. Ideally, the provenance flows should be continually linked to the subject data and should be recording any permutations.
- Is there an audit trail that clearly ties the digital evidence back to the initiating entry or data origin? Or, can this trail lead forward to the point of inclusion on the face of the financial statements? – Not offered in the core Hadoop platform, but this aspect of provenance may be added
- Does the Big Data electronic evidence include metadata that identifies who made the entry and when? – Core Hadoop does not record metadata outside of the storage and resource management points, but could be built into the Hadoop platform modifications.
- What are the controls designed to prevent unauthorized changes to the Big Data digital evidence after it was created? – The evidence of the enforcement

of these controls is available through access activity log files, which are minimally recorded in base Hadoop

- Who has or had access rights to change the Big Data digital evidence? – The evidence of enforcement of access rights is available only at two points in Hadoop
- How does the auditor know that the Big Data digital evidence hasn't been intentionally altered? –Core Hadoop can only provide metadata at two points
- Has the audit logging process been configured to record all access attempts, whether successful or not? – Core Hadoop is not configured for that degree of logging
- Have the audit logs been reviewed independently? - This control is independent of the Hadoop platform
- Has the continuity of logs been maintained and any gaps justified? – Core Hadoop does not provide enough metadata to determine this
- Have the logs been frequently copied to off-line, read only media and stored in a separate secure location, inaccessible to those who might be motivated to change it? – Hadoop, as originally configured, does not copy this information
- Has the access to the logs and their security settings been recorded, and limited to only authorized persons? – This information is not provided by core Hadoop and additional applications are required

Clearly, Hadoop requires applications that may contribute additional aspects of provenance to the basic platform (Lin and Ryaboy 2013). The section that follows proposes secure provenance recording, extended to the HadoopProv framework discussed earlier in this section.

4.6 The Big Data Provenance Black Box and Evidence Collection

4.6.1 The Big Data Provenance Black Box

All of the proposed systems to date make use of separate Big Data provenance storage files (Akoush et al 2013; Park and Lee 2013; Crawl et al 2011; Ikeda et al 2011; Park, Ikeda, and Widom 2011). However there is scant detail provided about a critical aspect of provenance for auditors: secure provenance record storage. Furthermore, these files are likely to be much larger than the Big Data files that they describe, as a many to one scenario (Ghoshal and Plale 2013; Buneman et al 2011). As such, the storage of provenance ought to be kept separate from the main files, so as to not encumber any processing overhead (Hasan, Sion and Winslett 2009). However, if the provenance is being frequently queried then there could be partial or full connections to the main workflow (Braun, Shinnar, and Seltzer 2008; Glavic 2014; Bao et al. 2009).

Storage of Big Data provenance files is as critical am aspect as the recording of the Big Data origins and transformations, since the storage should be secure (Hasan et al 2009). Maintaining the integrity and security of data provenance is further complicated by the fact that it is linked to the data itself. These linkages are also expressed as provenance and audit workflows. Basically, assurance needs to be provided that the provenance records of the data and the audit workflows themselves have not been altered or thwarted (Aldeco-Perez and Moreau 2010; Braun et al 2008) while being simultaneously connected to the Big Data itself.

This paper proposes a conceptual framework by which to achieve this secure storage of Big Data Provenance – that of a Big Data Provenance Black Box (BDPBB). The concept of a Black Box for provenance or log file storage is not a new concept and has been proposed previously (Stamatogiannakis et al 2015; Accorsi 2009; Alles et al, 2004; Oppliger and Rytz 2003). In fact, Oppliger and Rytz explain at length how digital signatures, although feasible for securing provenance information, should be deployed in digital black boxes to truly provide reliable and trustworthy evidence. This paper extends the concept of this digital black box to the issue of secure provenance tracking of Big Data in Hadoop, in support of reliable evidence collection for auditors.

Black Boxes on airplanes record cockpit conversations and sounds, as well as numerous digital measurements sent from many different sensors. The concept here is that everything is being recorded and stored in an orderly fashion, as separate logs of activities in case these actions need to be analyzed or audited in the future. Black Boxes may be regarded as a type of log recorder. Recording data provenance is basically creating logs of data about the activities of data point(s) or document (Glavic 2014; Ghoshal and Plale 2013; Muniswamy-Reddy, Macko and Seltzer 2010; Souiah, Francalanza, andSassone 2009). Expanding on an earlier work (Alles et al 2004) where Black Boxes were conceptualized as an internal audit tool and Black Box (BB) Log file, this paper proposes that such a BB concept would serve well in the capacity of a Big Data provenance collection system. The main difference with a Big Data

Provenance Black Box (BDPBB) and the BB log file is that the former is primarily concerned with all provenance data connected with a particular firm, whereas the latter is primarily interested in data pertaining to the audit of that firm (Alles et al 2004). The BDPPB would generate a much larger Big Data than it records, so it would be magnitudes larger than the data collected in the BB log file of Alles et al 2004. However, given the rapidly decreasing cost of data storage, it is possible that cost might be less of a prohibiting factor for the collection and storage of huge provenance files.

The BDPBB could record every transaction and alteration of the Big Data into the provenance files. It could also record less granular provenance or work flows, the level of which to be suggested by the auditor and undertaken by management.. The provenance data could be recorded in a standardized format, determined by and particular to each host and which would enable search algorithms to find certain data points at certain time recordings. This standard is necessary to avoid the BDPBB becoming a data dump, where finding anything would be prohibitive in effort and cost. No entry to the log could be altered after it is recorded; it would read-only. This read-only quality would make the BDPBB feasible for an audit trail (Bishop 2006). The provenance production would be write-once and the provenance query would be read-only.

The most important assurances for the BDPBB to provide are those of integrity, security, and confidentiality, as these qualities provide security (Braun et al 2008; Cheah and Plale 2012). The BDPBB has to maintain privacy and security with its contents as read-only. Furthermore, stringent access controls should be applied utilizing a role based approach (Ferraiolo and Kuhn 2009; Bishop 2006). Protecting the DPBB against tampering and alteration could be achieved with write once mediums. However, these mediums can be destroyed. Another possibility is to hand the BDPBB over to a trusted third party for protection (McDaniel et al 2010). However, this transfer would create its own set of security issues.

Or perhaps the firm could compute and transfer a digital signature of the BDPBB to this third party. After all, it is possible to detect if the BDPBB has been altered, by using digital signatures (Stamatogiannkis, Groth, and Boss 2015; Accorsi 2009; Hasan et al 2009; Tsai et al 2007; Bishop 2006; Alles et al 2004; Oppliger and Rytz 2003)). With digital signatures, a small change to the original data results in a huge difference to the hashed message (digital signature). It is computationally impossible to create two different documents that have the same digest; so if one document is altered, it would be impossible to create another document with the exact same digital signature. A digital signature would not reveal any information about the content of the BDPBB, only if the content has been altered (Accorsi 2009; Alles et al 2004). With digital signatures, not only is the storage of big data provenance records secure, but this security is assured. Furthermore, in a related study of secure Hadoop, the authors established that encryption and decryption measures only added about 5% overhead to MapReduce jobs (Park and Lee 2013).

The BDPBB would be made available to appropriate regulators and auditors; however even access and read, which are not active changes, will be recorded as part of the Big Data or document provenance. This BDPBB takes advantage of the digitization of the firm and the capacities of its ERP system, at little additional cost

(Park and Lee 2013). The provenance of the Big Data is maintained securely with the Black Box concept in a provenance enabled Hadoop platform, such as HadoopProv, as mentioned earlier in Section Four and shown in Figure 26 below. Provenance is captured at multiple points as indicated in both Map and Reduce, and is recorded at Map Prov File and Reduce Prov File.

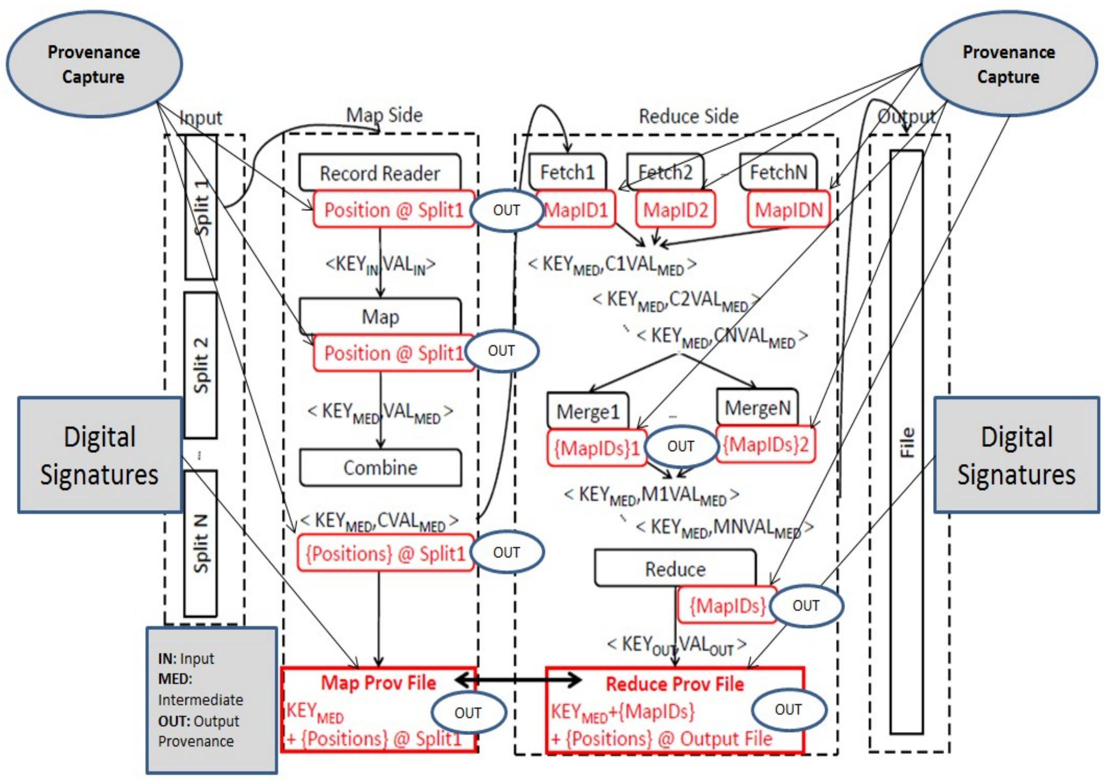

Figure 26: Big Data Provenance Black Box Illustration (modification of HadoopProv from Akoush et al, 2013)

The creators of HadoopProv suggested that the security of their provenance information is an area for future research (Akoush et al 2013). HadoopProv was conceived as an open source template which could be modified by others as needed. Provenance is captured and securely stored at the two separate phases of Map and

Reduce, with the secure provenance graph construction occurring later. This paper amends HadoopProv by suggesting that the provenance information be recorded as digital signatures and stored in a digital Black Box.

4.6.2 Evidence Collection with BDPBB and the Audit Standards Revisited

Businesses and their IT systems are becoming increasingly more complex and are constantly evolving, forcing the audit profession to constantly adjust examination processes. One such complexity is the use of external Big Data by clients to improve effectiveness and efficiency of business analytics. The auditor should regard external Big Data with increased professional skepticism. The BDPPBB may be regarded as one additional component in an integrated audit (ASB 2001, SAS 94), where the client is utilizing external Big Data and where the risk of insufficient competent evidence is greater. Thus, in the risk model of AR = IR x CR x DR, where audit risk (AR) is set low and inherent risk (IR) and control risk (CR) are assessed to calculate detection risk (DR), Big Data may significantly increase IR and CR. Detection Risk is the level of risk that the auditors could allow – high means that the auditor can afford less effective testing and low means the auditor will need more effective testing. Inherent Risk could be assessed high if the Big Data is external and the business process required substantial client judgement. CR could be high if the Big Data originated outside the client and was stored in the Cloud. For high risk IR and CR assertions and disclosures, the Big Data should be verified with fine grained provenance, with course provenance reserved for less risky areas. If the provenance does not exist or is not in BDPBB format, DR would be at a low level, see *Table 18*:

Data Type:	**Secure Provenance recorded or available?**	**Missing origins or steps?**	**Preliminary Detection Risk assessment of data type**
Paper external:	Yes	Yes	low/medium
		No	high
	No	Yes	low
		No	low/medium
Paper internal:	Yes	Yes	medium
		No	high
	No	Yes	low
		No	low/medium
Electronic external:	Yes	Yes	low/medium
		No	high
	No	Yes	low
		No	low/medium
Electronic internal:	Yes	Yes	medium
		No	high
	No	Yes	low
		No	low/medium
Big Data external:	Yes	Yes	low/medium
		No	high
	No	Yes	low
		No	low
Big Data internal:	Yes	Yes	Medium/low
		No	high
	No	Yes	low/medium

		No	low/medium

Table 18: Proposed DR assessment for each data type

In Figure Six, levels of DR for each data type are proposed, irrespective of whether the data is qualitative or quantitative. The lowest DR assessments for all data types exist when secure provenance does not exist for that data and there are gaps in either its origins other intermediate steps. Transactions or data types that are slightly less risky are indicated as low/medium and those that pose medium risk are highlighted in yellow. Data types that are high DR pose less risk of material misstatement to the auditor – the auditor, based on secure provenance of the data and its completeness should be able to afford less effective testing.

The high DR scenarios all assume that the client is recording fine and coarse provenance in a BDPBB format wherever and whenever external Big Data is acquired and that the client has agreed to secure and store this BDPBB outside its control for the benefit of auditors and regulators. Currently, this provenance recording may depend on the client's own assessment of its exposure to the risk of false information from external Big Data. However, businesses that have greater reliance on external Big Data may have a greater probability of being negatively impacted by faulty analyses derived from incompetent external Big Data.

Businesses such as Amazon, Twitter, Facebook, and several large banks and insurance companies have all experienced incidents due to faulty external Big Data social media and have responded with increased provenance collection efforts (Lin and Ryaboy 2013; Castillo et al 2011). Twitter can currently collect provenance on reads and writes but not the source control, due to the immensity of its data with its thousands of Hadoop nodes that process over 340 million tweets or 100 terabytes daily (Lin and Ryaboy 2013) – an improved provenance, but not complete. This lack of provenance origin is troublesome, as Twitter has disclosed that fraudulent accounts and tweet spam could diminish its platform (Twitter 2014). Furthermore, 10% of Twitter's revenue originates from data licensing, where data "partners" are allowed to access, search, and analyze public Tweets and their content (Twitter 2014).

However, as businesses rely more and more on external Big Data, it is hoped that the long term issues presented by the four V's (one of which is veracity or provenance) will be successfully be addressed by vendors, systems experts, and academics. Although businesses may be realizing short term benefits from acquiring and analyzing external Big Data, eventually the complexities presented by its four V's should be addressed. Secure provenance collection and storage of external Big Data will hopefully become standard processes.

The BDPBB would appear to be somewhat computationally expensive at this time, based on the studies of HadoopProv, secure Hadoop, and digital signatures. It would seem that the more provenance tracking to be collected as BDPBB and the more fine this provenance, the more expensive the process. The actual application of the BDPBB (or a similar platform) is an area for future case-study research regarding computational and monetary costs.

Section Two and Section Four reviewed how a skeptical auditor should regard electronic evidence, log files, and IT controls. These conditions can now be addressed again with the perspective of the proposed BDPBB:

- Is the Big Data electronic evidence subject to alteration without an audit trail or evidence of this change? – The audit trail is securely recorded in BDPBB, where any alteration that occurs with the subject data is recorded and this recording is write once, read only
- Is there an audit trail that clearly ties the Big Data digital evidence back to the initiating entry or data origin? Or, can this trail lead forward to the point of inclusion on the face of the financial statements? – With recording of provenance flows, this trail is available
- Does the Big Data electronic evidence include metadata that identifies who made the entry and when? – This metadata is now available from more points in the Hadoop process
- What are the controls designed to prevent unauthorized changes to the Big Data digital evidence after it was created? – Evidence of IC compliance is available through process logs that have been securely recorded in BDPBB
- Who has or had access rights to change the Big Data digital evidence? – Evidence of access rights compliance is available through additional metadata that is available in BDPBB
- How does the auditor know that the Big Data digital evidence hasn't been intentionally altered? This information is securely recorded in the BDPBB

- Has the audit logging process been configured to record all access attempts, whether successful or not? – This information is securely recorded in the BDPBB
- Have the audit logs been reviewed independently? – This control is maintained by limiting access to external auditors, internal auditors, and appropriate regulators
- Has the continuity of logs been maintained and any gaps justified? – Any changes to the provenance logs are securely maintained in the BDPBB.
- Have the logs been frequently copied to off-line, read only media and stored in a separate secure location, inaccessible to those who might be motivated to change it? – The provenance logs are continually updated as read-only and stored separately as digital signatures in a secure location with limited access
- Has the access to the logs and their security settings been recorded, and limited to only authorized persons? – The BDPBB records read-only information of all access attempts

Additionally, the audit standards specify attributes for reliable evidence, which may now be revisited in the context of the BDPBB (*Table 19*):

Evidence Characteristics:	Paper Evidence:	Electronic Evidence:	BDPBB Evidence
<u>Alterability:</u> easily altered evidence lacks credibility; evidence should be difficult to alter	Difficult to alter without detection	Alterations may be difficult to detect without performing specifically designed tests	Alterations of the data are easy to detect and verify with BDPBB files

Prima facie credibility: SAS 80 establishes a hierarchy of credibility –outside sources enhance credibility when independent of the client and confirmable	Outside sources of paper and documentary evidence and submitted directly to the auditor enhance credibility; inside sources of paper evidence that have been reviewed and processed by outsiders is also reliable	An electronic document derives its credibility primarily from the controls within the system. Outside electronic documentation/data is missing the assurance of system controls that the document or data is not fraudulent or altered	Outside sources are credible to the extent that their provenance has been securely recorded with the BDPBB. Auditors can readily determine the degree of veracity of the dataset based on its secure provenance
Completeness of documents: All essential terms of a transaction are verifiable	Typically all essential terms are included on its surface in a text/human readable form	An electronic system may substitute codes or cross-references to other data files that may not be accessible	The BDPBB file is complete in that it will show what has been altered and where the transaction evidence is incomplete
Evidence of approvals: This essential aspect of internal controls should be easily verifiable and transparent	Approvals integrated into paper documentation add to completeness	Electronic approvals may be similarly integrated, but need additional verification	BDPBB data can record the approvals as metadata/course grained provenance
Ease of use: Simplicity of application and access encourages compliance	Paper evidence can usually be evaluated without the use of additional tools and/or skills	Electronic evidence may require extraction of data by an expert	BDPBB could be designed with a simple interface for auditor interaction/query
Clarity: competent evidence should allow for the same re-performance and conclusions by other auditors	The nature of paper documentation is readily clear	The nature of electronic evidence is not always so clear, particularly in the absence of appropriate controls	BDPBB offers a straightforward recording of whether the provenance information has not been altered or not and the entire lineage of the dataset that is possible to record

Table 19: Evidence Characteristics of Paper, Electronic, and BDPBB format

Many of the concerns about audit evidence in electronic environments may be satisfied with secure provenance of the datasets. Metadata, log files, and provenance graphs can be recorded and stored securely for reference by the auditor regarding the evidence characteristics. Secure provenance enables to auditor to ascertain whether the data has been altered or not, or whether the origins of the data have been accounted for. Using provenance information, the auditor may more confidently and accurately assess the level of risk that the data poses to certain business accounting judgements, processes, and assumptions.

4.7 Discussion and Concluding Remarks

Big Data is now an important component of many businesses, due to the rapid development of social media, sensors, and IoT concurrent with increased data collection capabilities and storage capacity. Businesses, or audit clients, may be generating this Big Data internally or accessing it from external sources. Furthermore, this data has attributes of massive volume, high velocity, wide variety, and uncertain veracity (Zhang et al 2015). These four V's of Big Data persist as issues for entities attempting to unlock additional value from Big Data (CompTIA 2015). Basically, the Big Data trend may exhibit evidence of Amara's Law[9]: "the tendency to overestimate the effects of a technology in the short run and underestimate the effects in the long run". The Big Data attribute of uncertain veracity is particularly troubling, as this challenges the requirement of reliable competent audit evidence in the audit standards.

[9] Amara's Law is the statement that Dr. Roy Charles Amara, researcher and scientist, is well known for: "*We tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run.*" See https://en.wikipedia.org/wiki/Roy_Amara

Uncertain veracity in data means that the data lineage and transformations are not verifiable and not readily available. Lack of provenance in this instance equals unreliable data. Therefore, in a Big Data client environment, auditors may need to be more cognizant of secure data provenance.

The standards require that auditors ensure that the information generated through the client's system is reliable, before the audit opinion is generated (Li et al 2007; Alles et al 2002; Elliott 1997). This requirement of reliability verification exists regardless if the auditor examines few (sampling) or all transactions (continuous monitoring). Furthermore, Sarbanes-Oxley requires that auditors verify that the management report regarding Internal Controls is accurate, and such auditor attestation requires re-performance of transactions and controls. In an electronic environment, the only "map" of a transaction or data set may very well be the provenance record, also known as an audit trail. As Big Data increases in ubiquity of usage across businesses and industries, external auditors will be increasingly pressed to validate the reliability of this Big Data, particularly external Big Data and its attributes. This external Big Data may be messy, which clients may tolerate in the short term since the benefits of using the Big Data appear to outweigh the costs (Cukier and Mayer-Schoenberger 2013). However, SOX still requires management to provide auditable data, and auditors are not given the license, according to the current standards, to overlook the quality, reliability, and veracity of material audit evidence.

In this chapter, the BDPBB has been suggested as a possible means to provide secure data provenance of external Big Data that may serve as reliable audit evidence. Other solutions may exist that can address this issue – hopefully, this paper has stimulated more discussion about the secure provenance of Big Data for auditing. Basically, how should the truthfulness of the results of data analysis be validated by auditors when the data origins and/or permutations are unknown, as is often the case with MapReduce/Hadoop Big Data platforms? As such, without this provenance, Big Data which has been processed in MapReduce/Hadoop poses a huge risk as unreliable audit evidence when conducting audit examinations.

This chapter posed a conceptual model of a BDPBB based on HadoopProv, which has been demonstrated to be the most cost and work load efficient of any Hadoop provenance collection application to date (Alabi et al 2015; Akoush et al 2013). However, HadoopProv was not proposed as a secure system, and has been modified as the BDPBB here. Subsequent application and demonstration of the BDPBB in a Hadoop Big Data environment is an area for future research and exploration. Efficiency performance of a secure provenance system in Hadoop should be evaluated as should computational costs.

External Big Data that has been processed with Hadoop presents unique challenges of complexity and possibly high computational costs to the client and subsequently the auditing profession. In this context, to what extent should the auditing profession regard external Big Data as competent evidence and under what circumstances?

The Audit Standards should address the unique situation posed by Big Data: that external evidence in the form of external Big Data may not be reliable unless secure data provenance of that data has been recorded.

Finally, internal auditors may have more exposure than public auditors to examinations of business decisions and observations that were generated from "messy" external Big Data that was processed with Hadoop. In a survey by the Institute of Internal Auditors (IIA), nearly half of the auditors had little or no involvement with data quality evaluation, despite the fact that 23% of them had only slight or no confidence in that quality (Tysiac 2016). Perhaps the genesis of a solution that addresses the challenges of external Big Data audit evidence could occur initially within the internal auditing profession.

This chapter has contributed to the discussion of Issue 8 regarding secure data provenance in the Big Data environment, from a public auditing context. As businesses proceed to embrace Big Data and its potential for impactful and insightful analytics, this complex challenge of scant Big Data provenance and the subsequent erosion of evidence reliability should not be ignored by the audit profession, regulators, and academics.

CHAPTER FIVE: CONCLUDING COMMENTS

There is an increasing recognition in the public audit profession that the emergence of big data (Vasarhelyi, Kogan, and Tuttle 2015) as well as the growing use of analytics by audit clients has brought new concerns and opportunities. Financial auditing in the modern economy will soon require a paradigm change and this dissertation highlights some of the issues that need to be addressed for such a shift to occur. The first chapter introduces the following concerns:

1. What does previous research say about analytics in the audit engagement?
2. Should new (modern) analytics methods be used in the audit process?
3. Which of these methods are the most promising?
4. Where in the audit are these applicable?
5. Should auditing standards be changed to allow/facilitate these methods?
6. Should the auditor report be more informative?
7. What are the competencies needed by auditors in this environment?
8. How can the provenance of external Big Data provide assurance as audit evidence?

This dissertation contributes to the audit literature with its extensive elaboration of these issues and provides direction for future research. This research is relevant to:

- ✓ Audit academics and researchers who are interested in analytics and big data in the audit engagement,
- ✓ Practitioners or auditors who share these same concerns and are curious about the innovations in research about audit analytics, and

- ✓ Regulators who are seeking to update the standards and suggest best practices regarding the use of analytics in the engagement.

Before these many issues can be addressed, researchers should understand the scope of extant research. Keele (2007 p 3) states that “A systematic literature review…is a means of identifying, evaluating, and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest.” The second chapter addresses the first issue by means of the Systematic Literature Review Research Method (SLRRM): that is, what is the previous academic research? Additional concerns are also considered: should more complex analytics be used in the engagement and if so, where? Which techniques appear to be most promising? The audit standards provide minimal guidance. This chapter proposes that the answers to these questions may be assisted by an examination of the extant external audit research.

Before this study, a recent comprehensive synthesis of relevant published audit analytics research was not available. Accordingly, 301 papers are ultimately identified that discuss the use of analytical procedures in the public audit engagement. These papers are categorized by technique, engagement phase, and many other attributes for understanding. This analysis of the literature is categorized as an External Audit Analytics (EAA) framework, which is subsequently expanded with the concepts of business analytics (Holsapple et al, 2014). Specifically, this synthesis organizes the audit research, thereby offering guidelines regarding possible approaches for more complex and data driven analytics in the engagement.

The breadth and scope of approaches in the literature is astonishing, given the somewhat limited and narrow applications of analytics in practice. The fact that 301 papers discuss analytics in the audit engagement is significant. The enormity of the extant research is apparent and challenges the assumption that the profession has always been focused only on ratio analysis, sampling, and scanning. This literature review provides a significant contribution to the audit literature in that it:

- ✓ Summarizes and organizes the existing research about analytics and big data in the audit engagement
- ✓ Identifies gaps in this research by means of the EAA Framework
- ✓ Provides a framework/background – the EAA Framework – with which to understand this extant relevant research and to appropriately direct new research activities, practice, and regulations

This SLRRM shows that extant research has been undertaken not only regarding Audit Examination techniques, but also regarding regression, unsupervised, supervised, and other statistical approaches. This chapter details and then organizes[1] all the relevant research for any approach that occurs in a phase of the audit engagement. Academics, practitioners, and regulators may readily identify previous research for many techniques in the audit phases. For example, the PCAOB is re-assessing the feasibility of a more quantitative reporting format for the Audit Opinion and CAMs, as discussed at length in Chapter Three. Chapter Two provides an organized reference guide that directs attention to the papers that discuss various

[1] Table 20, Appendix A and Table 23, Appendix B

analytics and reporting techniques for that phase of the engagement. The PCAOB may see that 46 papers discuss analytics in the reporting phase, and these papers are identified[2]. Essentially, this second chapter directs the research process for the audit profession and exposes the degree of thought and analysis that has already occurred, thereby offering a significant contribution to the field.

It may be surprising that analytics in the engagement has been a widely debated topic for over fifty years, culminating in 301 papers. The challenge for academia is to help bridge the apparent chasm between this voluminous research, regulation, and practice. That is, the broad expanse of research regarding analytics in the engagement is now exposed, in juxtaposition to the very narrow range of analytics used by the external audit profession. What has been lacking to date is the execution in assurance practice of this rich research – however, with the challenges that auditors face in this modern business environment of analytics and big data, motivation for a shift in practice towards more complex analytics surely must be strengthening.

The third chapter elaborates on six additional major concerns facing the audit profession as business moves towards big data and advanced analytics, thereby contributing to the audit literature. These concerns are as follows:

1. Should new (modern) analytics methods be used in the audit process?
2. Which of these methods are the most promising?
3. Where in the audit are these applicable?
4. Should auditing standards be changed to allow/facilitate these methods?

[2] Table 20, Appendix A

5. Should the auditor report be more informative?
6. What are the competencies needed by auditors in this environment?

These issues are essential for the entire audit profession to examine if it is to successfully integrate more advanced analytics and big data in the engagement. Each issue is explored in detail, with implications for research and practice debated and recommendations organized in tables. Additionally, these issues may also help focus the research of concerns and gaps from the literature, as identified in the previous chapter. For example, academics may want to examine the literature available, or lack thereof, for the various techniques in the reporting phase and their contributions to the debate regarding increased quantitative disclosures in the opinion and Critical Audit Matters (CAM). Or, auditors may want to examine the literature about sampling in relation to 100% testing of populations,

Furthermore, additional research questions evolve from these six that seem to be also important to answer if EAA is to succeed in gaining widespread practical acceptance:

- ✓ How can analytics methods be used to create accurate expectation models for generating predictions to compare with actual accounting numbers? How should variable variances of predictions be chosen (Bumgarner and Vasarhelyi 2015)?
- ✓ What properties make a particular ADA technique more or less appropriate for a particular audit function?

- ✓ What types of "suspicion functions"[3] should be utilized in a preventive audit[4] in contrast to transaction or account reviews?
- ✓ How should the assurance function be reorganized to better accommodate analytics?
- ✓ How should the audit standards and processes be modified to enable and encourage the utilization of audit data analytics (ADA)?
- ✓ What would be the proper way of validating expectation models for ADA?
- ✓ What additional verification processes would be desirable with the extant analytic technology?
- ✓ Can the concept of "accuracy"[5] be defined for ADA? Should "accuracy" be defined by the standards? Is accuracy necessary to encourage the use of substantive audit analytics?

Additionally, a common thread of research questions relative to quantification are raised throughout this chapter and are elaborated upon here:

- ✓ Do modern disclosure and statistical methodologies make it possible, in certain cases, to automate pre-determined rules in order to perform procedures, derive results, and integrate these in a larger judgment?
- ✓ Can modern analytical methods be formalized regarding their applicability in different instances, their cumulative effects, and their classifications?

---

[3] A "suspicion function" is a linear multivariate equation that gives weights to characteristics of variables and analytical evidence to estimate its probability of being fallacious.

[4] Bungarner and Vasarhelyi (2015) decompose audit to retroactive and predictive approaches. A predictive audit may be preventative (when a suspicion score is large, a transaction is held for review), or just predictive to establish a standard for comparison.

[5] Acceptable relative error in engineering, equivalent to materiality in accounting

- ✓ If a midstream process detects a fault and activates an error correction process that is a mix of human judgment and automatic correction, is this an audit or control process? Does such a distinction make sense in the modern economy or are these differences becoming blurred?
- ✓ If a continuous layer detects "serious faults" (Vasarhelyi and Halper 1991) and stops a process, is this layer a part of operations, control, or audit?
- ✓ Can audit findings and judgments be disclosed in more a more disaggregate manner with the use of drill-down technologies where the opinion would be rendered and broken down into sub-opinions and quantified in terms of probabilistic estimates (Chesley 1975, 1976, 1977)?
- ✓ Would stochastic estimates in disclosures of Critical Audit Matters (CAM) be the more informative for the readers than deterministic statements that create illusory comfort? Should these expectations be reported to all stakeholders (e.g. investors, suppliers, analysts, etc.) or only to certain select parties?
- ✓ Should some of these exceptions be linked to smart contracts (Kosba et al. 2015) that automatically would execute a pre-agreed (e.g. covenant condition) action?

This third chapter contributes to research with its identification and discussions of the complexities resulting from applying big data and analytics in the engagement. Although many concerns are elaborated upon, it is possible that some are not mentioned since the scope of study in this area is rapidly expanding. As research and findings evolve in this domain, it is expected that some concerns will become less critical while others may unexpectedly gain urgency. However, what appears inevitable is the emerging overall importance that big data and analytics are posing to the audit profession, since they are dramatically changing the business environment and the capabilities of business processes. Business functions are changing, business capabilities are being added, anachronistic business functions are being eliminated, and processes are being substantially accelerated.

For example, one concern that is identified in Chapter One and mentioned in Chapter Three, that of the quality and reliability of big data, is gaining urgency for the audit profession as more businesses are integrated with the cloud, the Internet of Things (IoT), and exogenous data sources such as social media. The fourth chapter reflects this growing concern by focusing on the implications and additional considerations if big data is to be regarded as audit evidence, particularly that of external big data. The standards regard external sources of evidence as being highly reliable. However, in this age of big data many sources of evidence are untraceable and their origins unverifiable. The data may originate from sensors, videos, audio files, tweets, and other social media – all data types typically unfamiliar to the auditor (Warren et al. 2015). Basically, the questionable provenance of many sources of exogenous big data preclude it from being regarded as reliable audit evidence. Ideally, this big data should provide auditors the opportunity to apply more predictive and prescriptive analytics in the engagement (Holsapple et al. 2014), in addition to being regarded as extensive and reliable audit evidence. However, exogenous big data with questionable and insecure provenance cannot fulfill these roles for auditors. This chapter proposes a solution for providing secure provenance of big data, allowing it to be regarded as reliable evidence for external auditors.

For decision makers, researchers, auditors, and regulators, the ability to verify the accuracy information is of paramount importance. This chapter contributes to the literature by illuminating and discussing the challenge of provenance verification facing the auditor in the current big data information age. Furthermore, it identifies gaps in the audit and systems literature regarding secure big data provenance, and proposes a model and direction for future research – the Big Data Provenance Black Box (BDPBB).

However, although the BDPBB is illustrated as an efficient and effective means of secure provenance collection and storage, other solutions may exist or be developed that can address this issue. Additionally, to what extent should the auditing profession regard external big data as competent evidence and under what circumstances? The audit standards should assist the profession by providing clarification. As businesses proceed to embrace big data and its potential for impactful and insightful analytics, this complex challenge presented by sporadically available secure big data provenance should not be ignored by the audit profession, regulators, and academics. The illuminations contributed by this chapter present clear calls for research and investigation as to its feasibility and limitations.

5.1 Limitations

One possible limitation of this dissertation is that given the renewed and urgent interest in analytics and big data in the audit engagement, there may be very recent publications that are not included in this study. Another limitation may exist as well - there might be additional issues that have since been identified as being relevant to the profession regarding this topic. In short, given the recent expansive interest in this topic of big data and analytics in the external audit profession, there may be new papers and topics that are not covered. However, an online version of this study[6] will be updated periodically and will be available to interested researchers.

Another limitation exists regarding the BDPBB framework proposed in Chapter Four – its feasibility and efficiency are proposed here based on the earlier study results of its separate features but should be demonstrated in aggregate as the BDPBB in a case study setting. Furthermore, the arguments for and scope of BDPBB is based on the current audit standards - the BDPBB framework may need to be modified if there are adjustments to the regulations.

Another limitation exists in that this dissertation does not correlate the insights from the literature review of Chapter Two with the issues discussed in Chapters Three and Four. The literature review section serves to provide background for the research of each individual issue presented in Chapters Three and Four. The connections between the literature review and the issues that follow in Chapter Four are not clearly identified.

5.2 Calls for Future Research

This dissertation provides numerous opportunities for future research. From the literature review alone, each audit phase and technique should be examined more extensively. For example, based on the organization of the literature, further

[6] See Table 23, Appendix B

investigation is called for regarding ratio analysis, sampling, and artificial intelligence in the audit engagement. Or, additional research should be conducted about the absence of regression in most current engagement procedures. Or, the engagement and continuous activity phases of the engagement are sparse and/or absent of EAA, which should be addressed.

In short, the literature review organizes and analyzes the vast extant research and provides a framework of understanding based on business analytics, but does not provide further insights. However, it is significant and beneficial to the profession that the scope, details, and concentrations of the extant research are identified and organized. All gaps and areas of concentration identified in the External Audit Analytics (EAA) framework beg for further attention and research.

Finally, the topic of the viability of exogenous big data as audit evidence is called to question in Chapter Four and can only gain in importance for auditors since big data is gaining preference as the basis for business decisions and business analytics. Many predictive and prescriptive analytic techniques require big data to perform optimally – so the more that the auditor is required to rely on advanced EAA during an engagement, the more reliable, valid, and complete the data should be. It could be argued that the reliability, or the lack of proof of this big data is posing a major restraint on the audit profession's use of big data and subsequently, analytics, in the engagement. The use of analytics by auditors in the engagement may very well rest in the capability of research to address the main challenge presented by big data – can it be regarded as reliable audit evidence? The viability of big data as audit evidence must be addressed first before the standards and the profession can move forward using analytics that require the use of this big data. This dissertation calls attention to the audit profession that much additional research is required regarding the big data as audit evidence, if advanced analytics (EAA) are to be considered as appropriate engagement techniques.

This dissertation discusses and illuminates many issues facing the profession since businesses are becoming increasingly automated and are capturing massive amounts of data. Although businesses are embracing analytics and big data, the adoption of these innovations by external auditors has been restrained and cautious. These chapters contribute towards the research and development of solutions for many of these issues, and suggest areas for research that appear promising. These chapters lay the foundation for this future research by identifying and organizing the huge stream of literature in the audit profession regarding external audit analytics and reliability of big data as a source of audit evidence. It reviews the history of this research about analytics in the engagement and analyzes the components of this research, all towards conceptualizing a framework of EAA for the engagement. This dissertation also proposes where research should occur in the format of the EAA framework.

The issues of Big Data as audit evidence and the use (or lack thereof) of advanced EAA are intertwined, since the reliability of exogenous big data (which is often required for more advanced EAA), is an issue that must be resolved if EAA is to play a more substantive role in the engagement. Along with these complex discussions, these chapters raise a series of methodological and anticipatory questions as to how the public audit can be transformed by previous research into a modern audit. Finally, future research should elaborate on why it is necessary for external auditors to implement various EAA and what factors would encourage the augmentation of the engagement with big data and more complex EAA. It is anticipated that these chapters will assist with and encourage substantial future research and debate among academics, regulators, and the profession.

Analytics and big data have permeated business processes to the degree that most of the time the audit engagement occurs in a modern technical environment. Bernie Madoff's firm with its manually typed trade confirmations and customer account statements is an outlier practice of the past. Yet many of the audit standards and engagement practices that were established in the past specifically for the paper-driven audit persist to this day. The external audit profession must evolve if it is to keep pace with business practices and maintain its effectiveness and efficiency. It bears repeating: big data and analytics are dramatically changing the business environment and the capabilities of business processes. As a result, business functions are changing, business capabilities are being added, outdated business processes are being eliminated, and most of all, transactions and the amount of data describing them are substantially accelerating. The same must occur with the external audit function: its rules need to be changed, its steps evolved, automation integrated and augmenting its basic processes, and its timing should become almost instantaneous in predictive, prescriptive, and preventative analytical modes. Academics, regulators, and practitioners should avail themselves now of this dissertation with its vast literature review and numerous suggestions for research to address these urgent issues. The time has arrived for big data and analytics in the modern audit engagement.

REFERENCES:

Abdolmohammadi, M. J. (1986). Efficiency of the bayesian-approach in compliance testing-some empirical-evidence. *Auditing-A Journal of Practice & Theory*, *5(2)*, 1-16.

Accorsi, R., 2006. On the relationship of privacy and secure remote logging in dynamic systems. In *Security and privacy in dynamic environments* (pp. 329-339). Springer US.

Accorsi, R., 2009, September. Safe-keeping digital evidence with secure logging protocols: State of the art and challenges. In *IT Security Incident Management and IT Forensics, 2009. IMF'09. Fifth International Conference on* (pp. 94-110). IEEE.

Aggarwal, C. C. 2009. "Trio: A System for Data Uncertainty and Lineage." In *Managing and Mining Uncertain Data*, pp. 1-35.

Akoush, S., R. Sohan, and A. Hopper. 2013. "HadoopProv: towards provenance as a first class citizen in MapReduce." In *Proceedings of the 5th USENIX Workshop on the Theory and Practice of Provenance*, pp. 11-14. USENIX Association.

Alabi, O., Beckman, J., Dark, M. and Springer, J., 2015, June. Toward a Data Spillage Prevention Process in Hadoop using Data Provenance. In *Proceedings of the 2015 Workshop on Changing Landscapes in HPC Security* (pp. 9-13). ACM.

Aldeco-Pérez, R. and Moreau, L., 2010. Securing provenance-based audits. In *Provenance and Annotation of Data and Processes* (pp. 148-164). Springer Berlin Heidelberg

Allen, R. D., Beasley, M. S., & Branson, B. C. (1999). Improving analytical procedures: A case of using disaggregate multilocation data. *Auditing: A Journal of Practice & Theory, 18(2),* 128-142.

Alles, M. G. 2015. Drivers of the use and facilitators and obstacles of the evolution of Big Data by the audit profession. *Accounting Horizons* 29 (2): 439-449.

Alles, M. G., A. Kogan, and M. A. Vasarhelyi. 2004. "Restoring auditor credibility: tertiary monitoring and logging of continuous assurance systems." *International Journal of Accounting Information Systems* 5, no. 2: 183-202.

Alles, M. G., A. Kogan, and M.A. Vasarhelyi. 2008. Audit Automation for Implementing Continuous Auditing: Principles and Problems. Working paper, Rutgers University. Newark, N.J. USA.

Alles, M., G. Brennan, A. Kogan, and M.A. Vasarhelyi. 2006. "Continuous monitoring of business process controls: A pilot implementation of a continuous auditing system at Siemens".

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

Ameen, E. C., & Strawser, J. R. (1994). Investigating the use of analytical procedures: An update and extension. *Auditing: A Journal of Practice & Theory, 13(2),* 69.

American Institute of Certified Professional Accountants (AICPA). 2015. *Audit Analytics and Continuous Audit: Looking Toward the Future,* New York, NY

American Institute of Certified Public Accountants (AICPA). 1972. Responsibilities and Functions of the Independent Auditor. SAS No. 1, Section 110. New York, NY. Available at: http://www.aicpa.org/Research/Standards/AuditAttest/DownloadableDocuments/AU-00110.pdf

American Institute of Certified Public Accountants (AICPA). 1997. *Amendment to SAS no. 31, Evidential Matter.* SAS No. 80. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2012. *Audit Evidence*. Statement on Auditing Standards No. 122, AU-C Section 500. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2012a. Analytical Procedures. AU-C Section 520, Source: SAS No. 122. New York, NY. Available at: http://www.aicpa.org/Research/Standards/AuditAttest/DownloadableDocuments/AU-C-00520.pdf

American Institute of Certified Public Accountants (AICPA). 2015. *Continuous Auditing and Audit Analytics*. Monograph. New York, NY: AICPA.

American Institute of Certified Public Accountants. (1958). *Glossary of statistical terms for accountants and bibliography on the application of statistical methods to accounting, auditing and management control*. AICPA library, pp. 1-30.

American Institute of Certified Public Accountants. (AICPA). Audit Sampling Committee. 2012b. *Audit Sampling: Audit Guide*. New York, NY: AICPA.

Anderson, J. C., Kaplan, S. E., & Reckers, P. M. (1992). The effects of output interference on analytical procedures judgments. *Auditing: A Journal of Practice & Theory, 11(2),* 1.

Anderson, U., & Koonce, L. (1998). Evaluating the sufficiency of causes in audit analytical procedures. *Auditing: A Journal of Practice & Theory, 17(1),* 1.

Appelbaum, D. 2016. Securing Big Data Provenance for Auditors: The Big Data Provenance Black Box as Reliable Evidence. Working paper, Rutgers Business School, Newark N.J.

Appelbaum, D., Kogan, A., and Vasarhelyi, M.A. (2016). Continuous audit and big data with audit analytics: research needs. Working paper, Rutgers University. Newark, NJ

Appelbaum,D., D.S.Showalter, T. Sun, and M.A. Vasarhelyi. 2015. "Analytics Knowledge Required of a Modern CPA in this Real-Time Economy: A Normative Position". *Presented at the Accounting Information Systems Educator Conference, June 25-28, 2015.* Colorado Springs

Appelbaum,D., A. Kogan, and M. A. Vasarhelyi. 2016. “Analytics in External Auditing: A Literature Review”. Working paper, Rutgers Business School, Newark N.J. USA.

Arkin, H. (1957). Statistical sampling in auditing. *New York Certified Public Accountant (pre-1986)*, *27*(000007), 454.

Arkin, H. (1958). A statistician looks at accounting. *Journal of Accountancy (pre-1986), 105(000004), 66.*

Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I. and Zaharia, M., 2010. A view of cloud computing. *Communications of the ACM*, *53*(4), pp.50-58.

Arrington, C. E., Hillison, W., & Jensen, R. E. (1984). An application of analytical hierarchy process to model expert judgments on analytical review procedures. *Journal of Accounting Research*, 298-312.

Asare, S. K., & Wright, A. (1997). Hypothesis revision strategies in conducting analytical procedures. *Accounting, Organizations and Society*, *22*(8), 737-755.

Assunção, M.D., Calheiros, R.N., Bianchi, S., Netto, M.A. and Buyya, R., 2015. Big Data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, *79*, pp.3-15.

Atzori, L., A. Lera, and G. Morabito. 2010. The Internet of things: A survey. Computer Networks 54 (15): 2787–2805.

Auditing Standards Board (ASB), 1996. *Amendment to SAS 31, Evidential Matter.* Statement on Auditing Standards No. 80. New York, NY: ASB

Auditing Standards Board (ASB). 2001. *The Effect of Information Technology on the Auditor’s Consideration of Internal Control in a Financial Statement Audit.* Statement on Auditing Standards No. 94 (Amends *Statement on Auditing Standards No. 55, Consideration of Internal Control in a Financial Statement Audit*). New York, NY: ASB.

Ayata. 2012. http://ayata.com/the-eveolution-of-big-data-analytics/

Bailey, A. D., & Jensen, D. L. (1977). A Note on the Interface Between Compliance and Substantive Tests. *Journal of Accounting Research*, 293-299.

Banerjee, A., Bandyopadhyay, T., & Acharya, P. (2013). Data analytics: Hyped up aspirations or true potential. *Vikalpa*, *38(4),* 1-11.

Bao, Z., S. Cohen-Boulakia, S. B. Davidson, A. Eyal, and S. Khanna. 2009. "Differencing provenance in scientific workflows." In *Data Engineering, 2009. ICDE'09. IEEE 25th International Conference on*, pp. 808-819.

Basu, A. T. A. N. U. (2013). Five pillars of prescriptive analytics success. *Analytics Magazine*, 8-12.

Bates, A., Mood, B., Valafar, M. and Butler, K., 2013, February. Towards secure provenance-based access control in cloud environments. In*Proceedings of the third ACM conference on Data and application security and privacy* (pp. 277-284). ACM.

Bauer, S. and Schreckling, D., 2013. Data Provenance in the Internet of Things. In *EU Project COMPOSE, Conference Seminar*.

Beck, P. J., Solomon, I., & Tomassini, L. A. (1985). Subjective prior probability distributions and audit risk. *Journal of Accounting Research*, 37-56.

Beck, P.J. and Solomon, I. (1985). Ex Post Sampling Risks and Decision Rule Choice in Substantive Testing. *Auditing: A Journal of Practice & Theory* Vol. 4, No. 2 Spring 1985

Bedard, J. C. (1989). An archival investigation of audit program-planning. Auditing: A Journal of *Practice & Theory, 9(1),* 57-71.

Bedard, J. C., & Biggs, S. F. (1991). The effect of domain-specific experience on evaluation of management representations in analytical procedures. *Auditing: A Journal of Practice & Theory, 10,* 77-90

Bedard, J. C., & Graham, L. E. (2002). The effects of decision aid orientation on risk factor identification and audit test planning. *Auditing: A Journal of Practice & Theory, 21(2),* 39-56

Bedard, J. C., Deis, D. R., Curtis, M. B., & Jenkins, J. G. (2008). Risk monitoring and control in audit firms: A research synthesis. *Auditing: A Journal of Practice & Theory, 27(1),* 187-218

Bedard, J. C., Ettredge, M., & Johnstone, K. M. (2006). Using electronic audit workpaper systems in audit practice: Task analysis, learning, and resistance. *Learning, and Resistance (March 2006)*.

Bedard, J.C. and Biggs, S.F., 1991. Pattern recognition, hypotheses generation, and auditor performance in an analytical task. *Accounting Review,* pp.622-642.

Bedard, J.C., D.R. Deis, M.B. Curtis, and J.G. Jenkins. 2008. Risk Monitoring and control in audit firms: A research synthesis. *Auditing; A Journal of Practice & Theory,* Vol. 27, No. 1 (May), 187-218.

Bedingfield, J. P. (1975). The current state of statistical sampling and auditing. *Journal of Accountancy (pre-1986)*, *140*(000006), 48

Behn, B. K., Kaplan, S. E., & Krumwiede, K. R. (2001). Further evidence on the auditor's going-concern report: The influence of management plans. *Auditing: A Journal of Practice & Theory, 20(1),* 13-28

Bell, T. B., & Carcello, J. V. (2000). A decision aid for assessing the likelihood of fraudulent financial reporting. *Auditing: A Journal of Practice & Theory, 19(1),* 169-184

Bell, T. B., & Tabor, R. H. (1991). Empirical analysis of audit uncertainty qualifications. *Journal of Accounting Research*, 350-370.

Bell, T. B., Knechel, W. R., Payne, J. L., & Willingham, J. J. (1998). An empirical investigation of the relationship between the computerization of accounting systems and the incidence and size of audit differences. *Auditing: A Journal of Practice & Theory, 17(1),* 13.

Bell, T. B., Landsman, W. R., & Shackelford, D. A. (2001). Auditors' perceived business risk and audit fees: Analysis and evidence. *Journal of Accounting Research*, *39(1),* 35-43

Bell, T., & Knechel, W. R. (1994). Empirical analyses of errors discovered in audits of property and casualty insurers. *Auditing: A Journal of Practice and Theory*, *13(1),* 84

Bertsimas, D., & Kallus, N. (2014). From Predictive to Prescriptive Analytics. *arXiv preprint arXiv:1402.5481*.

Bhattacharjee, S., Kida, T., & Hanno, D. M. (1999). The impact of hypothesis set size on the time efficiency and accuracy of analytical review judgments. *Journal of Accounting Research*, 83-100.

Bible, L., Graham, L., & Rosman, A. (2005). The effect of electronic audit environments on performance. *Journal of Accounting, Auditing & Finance*, *20(1),* 27-42.

Biddle, G. C., Bruton, C. M., & Siegel, A. F. (1990). Computer-intensive methods in auditing: Bootstrap difference and ratio estimation. *Auditing: A Journal of Practice & Theory, 9(3).*

Bier, C., 2013, May. How usage control and provenance tracking get together-a data protection perspective. In *Security and Privacy Workshops (SPW), 2013 IEEE* (pp. 13-17). IEEE.

Bierstaker, J. L., Bedard, J. C., & Biggs, S. F. (1999). The role of problem representation shifts in auditor decision processes in analytical procedures. *Auditing: A Journal of Practice & Theory, 18(1),* 18-36.

Bierstaker, J. L., Brody, R. G., & Pacini, C. (2006). Accountants' perceptions regarding fraud detection and prevention methods. *Managerial Auditing Journal*, *21(5),* 520-535.

Biggs, S. F., & Wild, J. J. (1984). A note on the practice of analytical review. *Auditing: A Journal of Practice & Theory, 3(2),* 68-79.

Biggs, S. F., Messier, W. F., & Hansen, J. V. (1987). A descriptive analysis of computer audit specialists decision-making behavior in advanced computer environments. *Auditing: A Journal of Practice & Theory*, 6(2), 1-21.

Biggs, S. F., Mock, T. J., & Watkins, P. R. (1988). Auditor's use of analytical review in audit program design. *Accounting Review,* 148-161.

Biggs, S.F. and Wild, J.J., 1985. An investigation of auditor judgment in analytical review. *Accounting Review,* pp.607-633.

Biggs, S.F., Mock, T.J. and Simnett, R., 1999. Analytical procedures: Promise, problems and implications for practice. *Australian Accounting Review, 9(1),* p.42

Birnberg, J. G. (1964). Bayesian statistics: A review. *Journal of Accounting Research*, 108-116

Bishop, M. 2006. *Introduction to Computer Security.* Pearson Education

Blocher, E. and Bylinski, J.H., 1985. The influence of sample characteristics in sample evaluation. *Auditing: A Journal of Practice & Theory*, *5(1),* pp.79-90.

Blocher, E., & Cooper, J. C. (1988). A study of auditors analytical review performance. *Auditing: A Journal of Practice & Theory, 7(2),* 1-28.

Blocher, E., & Patterson Jr, G. F. (1996). The use of analytical procedures. *Journal of Accountancy*, *181(2),* 53.

Blocher, E., & Robertson, J. C. (1976). Bayesian sampling procedures for auditors: computer-assisted instruction. *Accounting Review*, 359-363.

Blocher, E., Esposito, R. S., & Willingham, J. J. (1983). Auditors' analytical review judgments for payroll expense. *Auditing: A Journal of Practice and Theory, 3(1),* 75-91.

Bollen, J., Mao, H. and Zeng, X., 2011. Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), pp.1-8.

Boritz, J. E., & Wensley, A. K. (1992). Evaluating expert systems with complex outputs: The case of audit planning. *Auditing: A Journal of Practice & Theory, 11(2),* 14.

Boritz, J. E., Zhang, P., & Aldersley, S. (1993). On Combining Evidence from Subpopulations into a Composite Conclusion*. *Contemporary Accounting Research*, *10(1),* 227-245.

Bose, R. and Frew, J., 2005. Lineage retrieval for scientific data processing: a survey. *ACM Computing Surveys (CSUR)*, *37*(1), pp.1-28.

Bratten, B., Gaynor, L. M., McDaniel, L., Montague, N. R., & Sierra, G. E. (2013). The audit of fair values and other estimates: The effects of underlying environmental, task, and auditor-specific factors. *Auditing: A Journal of Practice & Theory*, *32*(*sp1*), 7-44.

Braun, U., A. Shinnar, and M. I. Seltzer. 2008. "Securing Provenance." In *HotSecurity*.

Brown, B., Chui, M. and Manyika, J., 2011. Are you ready for the era of 'big data'. *McKinsey Quarterly*, *4*(2011), pp.24-35.

Brown-Liburd, H. and Vasarhelyi, M.A., 2015. Big Data and Audit Evidence. *Journal of Emerging Technologies in Accounting*, *12*(1), pp.1-16.

Brown-Liburd, H., and M.A. Vasarhelyi. 2015. Big Data and Audit Analytics. *Journal of Emerging technologies in Accounting*, Forthcoming.

Brown-Liburd, H., H. Issa, and D. Lombardi. 2015. Behavioral Implications of Big Data's Impact on Audit Judgment and Decision Making and Future Research Directions. *Accounting Horizons,* Vol. 29, No. 2. 2015, pp.451-468

Brynjolfsson, E., Hammerbacher, J. and Stevens, B., 2011. Competing through data: Three experts offer their game plans. *McKinsey Quarterly*, *4*, pp.36-47.

Budescu, D. V., Peecher, M. E., & Solomon, I. (2012). The joint influence of the extent and nature of audit evidence, materiality thresholds, and misstatement type on achieved audit risk. *Auditing: A Journal of Practice & Theory,* 31(2), 19-41

Bumgarner,N. and M.A Vasarhelyi. 2015. Auditing—A New View. *AUDIT ANALYTICS*, p.3.

Buneman, P. and Davidson, S.B., 2010. Data provenance--the foundation of data quality. *2015-03-14]. http://www. sei. cmu. edu/measurement/research/upload/Davidson. pdf*.

Buneman, P., and W.C. Tan. 2007. "Provenance in databases." In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, pp. 1171-1173.

Buneman, P., S. Khanna, and W.C. Tan. 2000. "Data provenance: Some basic issues." In *FST TCS 2000: Foundations of Software Technology and Theoretical Computer Science*, pp. 87-93. Springer Berlin Heidelberg.

Buneman, P., S. Khanna, and W.C. Tan. 2001."Why and where: A characterization of data provenance." In *Database Theory—ICDT 2001*, pp. 316-330. Springer Berlin Heidelberg.

Burgstahler, D. and Jiambalvo, J., 1986. Sample error characteristics and projection of error to audit populations. *Accounting Review*, pp.233-248.

Burgstahler, D., Glover, S. M., & Jiambalvo, J. (2000). Error projection and uncertainty in the evaluation of aggregate error. *Auditing: A Journal of Practice & Theory, 19(1),* 79-99.

Butler, S.A., 1985. Application of a decision aid in the judgmental evaluation of substantive test of details samples. *Journal of accounting Research,* pp.513-526.

Byrnes, P. 2014. Developing Automated Applications for Clustering and Outlier Detection: Data Mining Implications for Auditing Practice, PhD Dissertation, Rutgers Business School, Continuous Audit and reporting Lab, Newark, NJ, 2014

Calderon, T.G. and Cheh, J.J., 2002. A roadmap for future neural networks research in auditing and risk assessment. *International Journal of Accounting Information Systems, 3(4),* pp.203-236.

Cameron, G., 2003. Provenance and Pragmatics. In *Workshop on Data Provenance and Annotation, Edinburgh*.

Cao, M., Chychyla, R., & Stewart, T. (2015). Big Data analytics in financial statement audits. *Accounting Horizons*, *Vol 29, No. 2,* pp 423-429

Caramanis, C., & Lennox, C. (2008). Audit effort and earnings management. *Journal of accounting and economics*, *45(1),* 116-138.

Carcello, J. V., & Neal, T. L. (2000). Audit committee composition and auditor reporting. *The Accounting Review, 75(4),* 453-467.

Carcello, J. V., Hermanson, D. R., & Huss, H. F. (2000). Going-concern opinions: The effects of partner compensation plans and client size. *Auditing: A Journal of Practice & Theory, 19(1),* 67-77

Carey, P., & Simnett, R. (2006). Audit partner tenure and audit quality. *The Accounting Review, 81(3),* 653-676

Carson, E., Fargher, N. L., Geiger, M. A., Lennox, C. S., Raghunandan, K., & Willekens, M. (2012). Audit reporting for going-concern uncertainty: A research synthesis. *Auditing: A Journal of Practice & Theory*, *32(sp1),* 353-384.

Caster, P. and Verardo, D., 2007. Technology changes the form and competence of audit evidence. *The CPA Journal*, *77*(1), p.68.

Castillo, C., Mendoza, M. and Poblete, B., 2011, March. Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web* (pp. 675-684). ACM.

Cerullo, M.V. and Cerullo, M.J., 2003. Impact of sas no. 94 on computer audit techniques. *Information Systems Control Journal*, *1*, pp.53-58.

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

Chambers, A., 2014. New guidance on internal audit–an analysis and appraisal of recent developments. *Managerial Auditing Journal*, *29*(2), pp.196-218.

Chambers, A., 2014. New guidance on internal audit–an analysis and appraisal of recent developments. *Managerial Auditing Journal*, *29*(2), pp.196-218.

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

Chang, A. M., Bailey Jr, A. D., & Whinston, A. B. (1993). Multi-auditor decision making on internal control system reliability: A default reasoning approach. *Auditing: A Journal of Theory & Practice, 12(2),* 1

Che, D., Safran, M. and Peng, Z., 2013, April. From big data to big data mining: challenges, issues, and opportunities. In *Database Systems for Advanced Applications* (pp. 1-15). Springer Berlin Heidelberg.

Cheah, Y.W., and B. Plale. 2012. "Provenance analysis: Towards quality provenance." In *E-Science (e-Science), 2012 IEEE 8th International Conference on*, pp. 1-8.

Chen, K. C., & Church, B. K. (1992). Default on debt obligations and the issuance of going-concern opinions. *Auditing: A Journal of Theory & Practice, 11(2),* 30.

Chen, P. and Plale, B.A., 2015, May. Big Data Provenance Analysis and Visualization. In *Cluster, Cloud and Grid Computing (CCGrid), 2015 15th IEEE/ACM International Symposium on* (pp. 797-800). IEEE

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

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

Cheney, J., L. Chiticariu, and W. C. Tan. 2009. *Provenance in databases: Why, how, and where*. Vol. 1, no. 4. Now Publishers Inc.

Chesley, G. R. (1975). Elicitation of subjective probabilities: a review. *Accounting Review*, 325-337.

Chesley, G. R. (1977). Subjective Probability Elicitation: The Effect of Congruity of Datum and Response Mode on Performance. *Journal of Accounting Research*, 1-11

Chesley, G. R. (1978). Subjective probability elicitation techniques: A performance comparison. *Journal of Accounting Research*, 225-241

Christensen, B. E., Elder, R. J., & Glover, S. M. (2014). Behind the Numbers: Insights into Large Audit Firm Sampling Policies. *Accounting Horizons*, *29*(1), 61-81

Christensen, B. E., Glover, S. M., & Wood, D. A. (2012). Extreme estimation uncertainty in fair value estimates: Implications for audit assurance. *Auditing: A Journal of Practice & Theory, 31(1),* 127-146

Christensen, C. 2013. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Review Press

Christensen, C. 2013. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Review Press.

Chu, Z., Gianvecchio, S., Wang, H. and Jajodia, S., 2012. Detecting automation of twitter accounts: Are you a human, bot, or cyborg?. *Dependable and Secure Computing, IEEE Transactions on*, *9*(6), pp.811-824.

Church, B. K., J.J.McMillan and A. Schneider. 2001. Factors affecting internal auditors' consideration of fraudulent financial reporting during analytical procedures. *Auditing: A Journal of Practice & Theory*, 20(1), 65-80

Cianci, A. M., & Bierstaker, J. L. (2009). The impact of positive and negative mood on the hypothesis generation and ethical judgments of auditors. *Auditing: A Journal of Practice & Theory, 28(2),* 119-144

Cleary, R., & Thibodeau, J. C. (2005). Applying digital analysis using Benford's law to detect fraud: the dangers of type I errors. *Auditing: A Journal of Practice & Theory, 24(1),* 77-81

Coakley, J.R., 1995. Using pattern analysis methods to supplement attention directing analytical procedures. *Expert Systems with Applications*, *9*(4), pp.513-528.

Coglitore, F., & Berryman, R. G. (1988). Analytical Procedures: A Defensive Necessity. *Auditing: A Journal of Practice & Theory, 7(2*), 150-163

Cohen, J.C. and Acharya, S., 2014. Towards a trusted HDFS storage platform: Mitigating threats to Hadoop infrastructures using hardware-accelerated encryption with TPM-rooted key protection. *Journal of Information Security and Applications*, *19*(3), pp.224-244

Cohen-Boulakia, S., O. Biton, S. Cohen, and S. Davidson. 2008. "Addressing the provenance challenge using ZOOM." *Concurrency and Computation: Practice and Experience* 20, no. 5: 497-506.

Colbert, J.L. and Smalling, J., 2011. EDI And The Financial Auditor. *Review of Business Information Systems (RBIS)*, *2*(4), pp.9-16.

Colbert, J.L., 1991. Statistical or non-statistical sampling: Which approach is best?. *Journal of Applied Business Research (JABR), 7(2),* pp.117-120.

Colon, R. (2015). Independent Auditors' Responsibilties for Violations of Anti-bribery Provisions Under the U.S. Foreign Corrupt Practices Act: Auditing for Bribes. *Journal of Forensic and Investigative Accounting,* Vol. 7, Issue 2, July - December 2015

CompTIA. 2015. "Big Data Insights and Opportunities." Research Report, November 2015. https://www.comptia.org/resources/big-data-insights-and-opportunities-2015

Corless, J. C. (1972). Assessing prior distributions for applying Bayesian statistics in auditing. *Accounting Review*, 556-566

Crawl, D., J. Wang, and I. Altintas. 2011. "Provenance for MapReduce-based data-intensive workflows." In *Proceedings of the 6th workshop on Workflows in support of large-scale science*, pp. 21-30.

Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A. and Tesconi, M., 2015. Fame for sale: efficient detection of fake Twitter followers. *Decision Support Systems*, *80*, pp.56-71.

Crosby, M. A. (1980). Implications of prior probability elicitation on auditor sample size decisions. *Journal of Accounting Research*, 585-593.

Crosby, M. A. (1981). Bayesian statistics in auditing: a comparison of probability elicitation techniques. Accounting Review, 355-365

Crosby, M. A. (1985). The development of Bayesian decision-theoretic concepts in attribute sampling. *Auditing-A Journal of Practice & Theory, 4(2),* 118-132

Cui, Y., and J. Widom. 2003. "Lineage tracing for general data warehouse transformations." *The VLDB Journal—The International Journal on Very Large Data Bases* 12, no. 1: 41-58

Cukier, K. and Mayer-Schoenberger, V., 2013. Rise of Big Data: How it's Changing the Way We Think about the World, The. *Foreign Aff.*, *92*, p.28

Cukier,K., and V. Mayer-Schoenberger. 2013. Rise of Big Data: How it's Changing the Way We Think about the World, The. *Foreign Affairs*, *92*, p.28

Cushing, B. E. (1974). A mathematical approach to the analysis and design of internal control systems. *Accounting Review*, 24-41.

Cushing, B. E., & Loebbecke, J. K. (1986). *Comparison of audit methodologies of large accounting firms*. Studies in Accounting research #26, American Accounting Association

Cushing, B. E., and Loebbecke, J.K. (1983). Analytical Approaches to Audit Risk: A Survey and Analysis. *Auditing: A Journal of Practice & Theory* Vol. 3, No. 1 Fall 1983

da Silva, C. G., & Carreira, P. M. (2012). Selecting Audit Samples Using Benford's Law. *Auditing: A Journal of Practice & Theory, 32(2),* 53-65

Dai,J., and M.A. Vasarhelyi. 2016. Audit 4.0. *Journal of Emerging Technologies in Accounting*. 2016.

Daroca, F. P., & Holder, W. W. (1985). The use of analytical procedures in review and audit engagements. *Auditing-A Journal of Practice & Theory, 4(2),* 80-92.

Daroca,F. P., and W.W. Holder. 1985. The use of analytical procedures in review and audit engagements. *Auditing-A Journal of Practice & Theory*, 4(2), 80-92

Davenport, T. H., and J. Kim. 2013. Keeping Up with the Quants. *Harvard Business Review Press,* USA.

Davenport, T. H., and L. K. Johnson. 2008. Competing on Analytics: The New Science of Winning. *Harvard Business School Press.*

Davidson, S. B., and J. Freire. 2008. "Provenance and scientific workflows: challenges and opportunities." In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pp. 1345-1350.

Davidson, S. B., S. Cohen-Boulakia, A. Eyal, B. Ludäscher, T. M. McPhillips, S. Bowers, M. Kumar Anand, and J. Freire. 2007. "Provenance in Scientific Workflow Systems." *IEEE Data Eng. Bull.* 30, no. 4 : 44-50.

Davis, G. B., & Weber, R. (1986). The Impact Of Advanced Computer-Systems On Controls And Audit Procedures-A Theory And An Empirical-Test. *Auditing: A Journal of Practice & Theory, 5(2),* 35-49.

Deakin, E. B., & Granof, M. H. (1974). Regression analysis as a means of determining audit sample size. *Accounting Review*, 764-771

Deakin, E.B., 1976. Distributions of financial accounting ratios: some empirical evidence. *The Accounting Review, 51(1),* pp.90-96.

Dean, J., and S. Ghemawat. 2008. "MapReduce: simplified data processing on large clusters." *Communications of the ACM* 51, no. 1: 107-113

Debreceny, R., Lee, S. L., Neo, W., & Shuling Toh, J. (2005). Employing generalized audit software in the financial services sector: Challenges and opportunities. *Managerial Auditing Journal*, *20*(6), 605-618

Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting Material Accounting Misstatements*. *Contemporary accounting research*, *28*(1), 17-82.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the sec*. *Contemporary accounting research*, *13*(1), 1-36.

Delen, D., & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359-363.

Delen,D., and H. Demirkan. 2013. Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359-363

Deshmukh, A. (1999). The role of audit technology and extension of audit procedures in strategic auditing. *International Journal of Applied Quality Management*, *2*(2), 187-209

Dilla, W., Janvrin, D.J. and Raschke, R., 2010. Interactive data visualization: New directions for accounting information systems research. *Journal of Information Systems*, *24*(2), pp.1-37.

Dohrer, R., M.A. Vasarhelyi, and P.McCollough. 2015. Audit Data Analytics. Presentation delivered to the IAASB, September 23.

Domingos, P. 2012. A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78-87

Dopuch, N., Holthausen, R. W., & Leftwich, R. W. (1987). Predicting audit qualifications with financial and market variables. *Accounting Review*, 431-454.

Dugan, M. T., Gentry, J. A., & Shriver, K. A. (1985). The X-11 Model-A New Analytical Review Technique For The Auditor. *Auditing-A Journal of Practice & Theory, 4(2),* 11-22

Duncan, M., 2015. Future Casting Influence Capability in Online Social Networks.

Durney, M., Elder, R. J., & Glover, S. M. (2013). Field data on accounting error rates and audit sampling. *Auditing: A Journal of Practice & Theory, 33(2),* 79-110.

Durney, M.T., Elder, R.J. and Glover, S.M., 2013. Error rates, error projection, and consideration of sampling risk: Audit sampling data from the field. Error Projection, and Consideration of Sampling Risk: Audit Sampling Data from the Field (April 5, 2013).

Durtschi, C., Hillison, W., & Pacini, C. (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. *Journal of forensic accounting, 5(1),* 17-34.

Dusenbury, R. B., Reimers, J. L., & Wheeler, S. W. (2000). The audit risk model: An empirical test for conditional dependencies among assessed component risks. *Auditing: A Journal of Practice & Theory, 19(2),* 105-117

Dusenbury, R., Reimers, J. L., & Wheeler, S. (1996). An empirical study of belief-based and probability-based specifications of audit risk. *Auditing: A Journal of Practice & Theory, 15(2),* 12

Dutta, S. K. & Srivastava, R. P (1993). Aggregation of evidence in auditing: A likelihood perspective. *Auditing: A Journal of Practice & Theory, 12 Supplement*

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

Dzuranin, A. C., and I. Malaescu. 2016. The Current State and Future Direction of IT Audit: Challenges and Opportunities. *Journal of Information Systems.* Vol. 30, No. 1. Spring 2016, pp. 7-20.

Edwards, W. (1995). Number magic, auditing acid and materiality-A challenge for auditing research. *Auditing: A Journal of Practice & Theory, 14(2),* 176

Eilifsen, A., & Messier Jr, W. F. (2014). Materiality guidance of the major public accounting firms. *Auditing: A Journal of Practice & Theory*, *34*(2), pp.3-26.

Eining, M. M., Jones, D. R., & Loebbecke, J. K. (1997). Reliance on decision aids: An examination of auditors' assessment of management fraud. *Auditing: A Journal of Practice & Theory, 16(2)*

Elder, R. J., A.D. Akresh, S.M. Glover, J.L. Higgs, and J. Liljegren. 2013. Audit sampling research: A synthesis and implications for future research. *Auditing: A Journal of Practice & Theory*, *32*(sp1), 99-129.

Elder, R. J., Akresh, A. D., Glover, S. M., Higgs, J. L., & Liljegren, J. (2013). Audit sampling research: A synthesis and implications for future research. *Auditing: A Journal of Practice & Theory*, *32*(sp1), 99-129.

Elder, R.J. and Allen, R.D., 1998. An empirical investigation of the auditor's decision to project errors. *Auditing: A Journal of Practice & Theory, 17(2),* p.71.

Elder, R.J. and Allen, R.D., 2003. A longitudinal field investigation of auditor risk assessments and sample size decisions. *The Accounting Review, 78(4),* pp.983-1002.

Elliott, R. K., & Rogers, J. R. (1972). Relating statistical sampling to audit objectives. *Journal of Accountancy*, 134.

Elliott, R.K. (1983). Unique Audit Methods: Peat Marwick International. *Auditing: A Journal of Practice & Theory Vol. 2, No. 2 Spring 1983*

Elliott, R.K., 1997. Assurance service opportunities: Implications for academia. *Accounting Horizons*, *11*(4), p.61.

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

Entwistle, G. and Lindsay, D., 1994. An archival study of the existence, cause, and discovery of. *Contemporary Accounting Research, 11(1),* p.271.

Erdogan, N. and Uludag, S., 2014. Comparison of Analysis Performed by Classical Approach and Bayesian Approach in Auditors' Decision Making Process. *Procedia-Social and Behavioral Sciences, 150,* pp.668-677.

Etheridge, H.L., Sriram, R.S. and Hsu, H.Y., 2000. A comparison of selected artificial neural networks that help auditors evaluate client financial viability. *Decision Sciences, 31(2),* pp.531-550.

Evans, J. R., and C. H. Lindner. 2012. Business Analytics: The Next Frontier for Decision Sciences. *Decision Line,* 43 (2), pp. 4-6.

Fanning, K. M., & Cogger, K. O. (1998). Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems in Accounting, Finance & Management, 7(1),* 21-41.

Felix Jr, W. L., & Kinney Jr, W. R. (1982). Research in the auditor's opinion formulation process: State of the art. *Accounting Review,* 245-271..

Felix, W. L. (1976). Evidence on alternative means of assessing prior probability distributions for audit decision making. *Accounting Review*, 800-807

Felix, W. L., & Grimlund, R. A. (1977). A sampling model for audit tests of composite accounts. *Journal of Accounting Research*, 23-41

Felix, W. L., & Niles, M. S. (1988). Research in internal control evaluation. *Auditing-A Journal of Practice & Theory, 7(2),* 43-60

Fellingham, J. C., Newman, D. P., & Patterson, E. R. (1989). Sampling information in strategic audit settings. *Auditing-A Journal of Practice & Theory, 8(2),* 1-21

Ferraiolo, D. F., and D. R. Kuhn. 2009. "Role-based access controls." *arXiv preprint arXiv:0903.2171*

Finley, D. R. (1983). Normal Form Decision Theory Development of the Audit Sampling Model. *Auditing: A Journal of Practice & Theory, Vol. 3, No. 1 Fall 1983*

Finley, D. R., & Boockholdt, J. L. (1987). A Continuous Constrained Optimization Model for Audit Sampling. *Auditing-A Journal of Practice & Theory, 6(2),* 22-39

Fischer, M. J. (1996). “Realizing” the benefits of new technologies as a source of audit evidence: An interpretive field study. *Accounting, Organizations and Society, 21(2),* 219-242.

Fraser, I.A., Hatherly, D.J. and Lin, K.Z., 1997. An Empirical Investigationof The Use Of Analytical Review By External Auditors. *The British Accounting Review, 29(1),* pp.35-47.

Freeman, S. 2015. Special report: Engaging lines of defense. *Freeman*, *31*(4), 21.

Freire, J., D. Koop, E. Santos, and C. T. Silva. 2008. "Provenance for computational tasks: A survey." *Computing in Science & Engineering* 10, no. 3: 11-21.

Frew, J., D. Metzger, and P. Slaughter. 2008. "Automatic capture and reconstruction of computational provenance." *Concurrency and Computation: Practice and Experience* 20, no. 5: 485-496.

Fukukawa, H., & Mock, T. J. (2011). Audit risk assessments using belief versus probability. *Auditing: A Journal of Practice & Theory, 30(1),* 75-99.

Gaeremynck, A., & Willekens, M. (2003). The endogenous relationship between audit-report type and business termination: Evidence on private firms in a non-litigious environment. *Accounting and Business Research, 33(1),* 65-79.

Galhardas, H., Florescu, D., Shasha, D., Simon, E. and Saita, C.A., 2001, June. Improving Data Cleaning Quality Using a Data Lineage Facility. In *DMDW* (p. 3).

Geerts, G. L., L. E. Graham, E. G. Mauldin, W. E. McCarthy, and V. J. Richardson. 2013. Integrating information technology into accounting and practice. *Accounting Horizons.* 27 (4): 815-840.

Geiger, M. A., & Raghunandan, K. (2001). Bankruptcies, audit reports, and the Reform Act. *Auditing: A Journal of Practice & Theory, 20(1),* 187-195

Ghoshal, D., and B. Plale. 2013. "Provenance from log files: a BigData problem." In *Proceedings of the Joint EDBT/ICDT 2013 Workshops*, pp. 290-297.

Gillett, P. R., & Srivastava, R. P. (2000). Attribute sampling: A belief-function approach to statistical audit evidence. *Auditing: A Journal of Practice & Theory, 19(1),* 145-155

Glancy, F. H., & Yadav, S. B. (2011). A computational model for financial reporting fraud detection. *Decision Support Systems, 50(3*), 595-601

Glavic, B. 2014. "Big Data Provenance: Challenges and Implications for Benchmarking." In *Specifying Big Data Benchmarks*, pp. 72-80.

Glavic, B., and K. R. Dittrich. 2007. "Data Provenance: A Categorization of Existing Approaches." In *BTW*, vol. 7, no. 12, pp. 227-241.

Glover, S. M., Jiambalvo, J., & Kennedy, J. (2000). Analytical procedures and audit-planning decisions. *Auditing: A Journal of Practice & Theory, 19(2),* 27-45

Glover, S. M., Prawitt, D. F., & Drake, M. S. (2014). Between a Rock and a Hard Place: A Path Forward for Using Substantive Analytical Procedures in Auditing Large P&L Accounts Commentary and Analysis. *Auditing: A Journal of Practice and Theory*. *34(3),* pp.161-179.

Glover, S. M., Prawitt, D. F., & Wilks, T. J. (2005). Why do auditors over-rely on weak analytical procedures? The role of outcome and precision. *Auditing: A Journal of Practice & Theory, 24(s-1),* 197-220

Glover, S.M., Prawitt, D.F. and Drake, M.S., 2014. Between a Rock and a Hard Place: A Path Forward for Using Substantive Analytical Procedures in Auditing Large P&L Accounts: Commentary and Analysis. *Auditing: A Journal of Practice & Theory*, *34*(3), pp.161-179.

Goble, C., 2002, October. Position statement: Musings on provenance, workflow and (semantic web) annotations for bioinformatics. In *Workshop on Data Derivation and Provenance, Chicago* (Vol. 3).

Godfrey, J. T., & Andrews, R. W. (1982). A finite population bayesian model for compliance testing. *Journal of Accounting Research,* 304-315

Graham, L.E., Damens, J. and Van Ness, G. (1991). Developing Risk Advisor: An Expert System for Risk Identification. *Auditing-a Journal of Practice & Theory, 10(1),* 69-96

Green, B. P., & Choi, J. H. (1997). Assessing the risk of management fraud through neural network technology. *Auditing: A Journal of Practice & Theory, 16(1),* 14

Green, W. (2008). Are Industry Specialists More Efficient and Effective in Performing Analytical Procedures? A Multi-stage Analysis. *International Journal of Auditing, 12(3),* 243-260.

Grimlund, R. A. (1988). Sample-Size Planning For The Moment Method Of Mus-Incorporating Audit Judgments. *Auditing-A Journal of Practice & Theory, 7(2),* 77-104

Grobstein, M., & Craig, P. W. (1984). A risk analysis approach to auditing. *Auditing-A Journal of Practice & Theory, 3(2), 1-16*

Hall, T. W., Pierce, B. J., & Ross, W. R. (1989). Planning Sample Sizes for Stringer-Method Monetary Unit and Single-Stage Attribute Sampling Plans. *Auditing-A Journal of Practice & Theory, 8(2), 64-89*

Ham, J., Losell, D. and Smieliauskas, W., 1985. An empirical study of error characteristics in accounting populations. *Accounting Review,* pp.387-406.

Hammerbacher, J., 2009. Information platforms and the rise of the data scientist. *Beautiful Data: The Stories Behind Elegant Data Solutions.* O'Reilly Media, pp.73-84.

Hammersley, J. S. (2011). A review and model of auditor judgments in fraud-related planning tasks. *Auditing: A Journal of Practice & Theory, 30(4),* 101-128.

Hammersley, J. S., Johnstone, K. M., & Kadous, K. (2011). How do audit seniors respond to heightened fraud risk?. *Auditing: A Journal of Practice & Theory, 30(3),* 81-101

Hand, J. R. (2005). The value relevance of financial statements in the venture capital market. *The Accounting Review, 80(2),* 613-648

Hansen, J. V., & Messier, W. F. (1986). A preliminary investigation of EDP-XPERT. *Auditing-A Journal of Practice & Theory, 6(1), 109-123*

Hardy, C.A. and Laslett, G., 2014. Continuous Auditing and Monitoring in Practice: Lessons from Metcash's Business Assurance Group. *Journal of Information Systems*, *29*(2), pp.183-194.

Harper, R. M., Strawser, J. R., & Tang, K. (1990). Establishing investigation thresholds for preliminary analytical procedures. *Auditing-A Journal of Practice & Theory, 9(3),* 115-133

Hasan, R., R. Sion, and M. Winslett. 2009. "The Case of the Fake Picasso: Preventing History Forgery with Secure Provenance." In *FAST*, vol. 9, pp. 1-14.

Heiman, V.B., 1990. Auditors' assessments of the likelihood of error explanations in analytical review. *Accounting Review*, pp.875-890.

Hill, H. P. (1958). An accountant looks at statistics. *Journal of Accountancy* (pre-1986), 105(000004), 57

Hirst, E.D. and L. Koonce. (1996). Audit Analytical Procedures: A Field Investigation. *Contemporary Accounting Research Vol. 13 No. 2 (Fall 1996)* pp. 457-486

Hitzig, N. B. (1998). Detecting and estimating misstatement in two-step sequential sampling with probability proportional to size. *Auditing: A Journal of Practice & Research, 17(1),* 54

Hogan, C. E., Rezaee, Z., Riley Jr, R. A., & Velury, U. K. (2008). Financial statement fraud: Insights from the academic literature. *Auditing: A Journal of Practice & Theory, 27(2),* 231-252

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

Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. *Decision Support Systems, 64,* 130-141

Holsapple,C., A.Lee-Post, and R. Pakath. 2014. A unified foundation for business analytics. *Decision Support Systems*, 64: 130-141

Hoogduin, L. A., Hall, T. W., & Tsay, J. J. (2010). Modified sieve sampling: A method for single-and multi-stage probability-proportional-to-size sampling. *Auditing: A Journal of Practice & Theory, 29(1),* 125-148

Hopwood, W., McKEOWN, J. C., & Mutchler, J. F. (1994). A Reexamination of Auditor versus Model Accuracy within the Context of the Going-Concern Opinion Decision*. *Contemporary Accounting Research, 10(2),* 409-431

Horgan, J. M. (1997). Stabilising the sieve sample size using PPS. *Auditing: A Journal of Practice & Theory, 16(2),* 40

Houghton, C. W., & Fogarty, J. A. (1991). Inherent risk. *Auditing-a Journal of Practice & Theory, 10(1),* 1-21

Hylas, R. E., & Ashton, R. H. (1982). Audit detection of financial statement errors. *Accounting Review,* 751-765

IBM. 2013. Descriptive, predictive, prescriptive: Transforming asset and facilities management with analytics. *Thought Leadership White Paper*, October 2013.

Ijiri, Y. and Leitch, R.A., 1980. Stein's Paradox and Audit Sampling. *Journal of Accounting Research*, pp.91-108

Ijiri, Y., & Kaplan, R. S. (1969). The Four Roles of Sampling in Auditing: Representative, Corrective, Protective, and Preventive (No. RR-165). *Carnegie-Mellon Univ Pittsburgh Pa Management Sciences Research Group*.

Ijiri, Y., & Kaplan, R. S. (1971). A model for integrating sampling objectives in auditing. *Journal of Accounting Research*, 73-87.

Ismail, Z., & Trotman, K. T. (1995). The impact of the review process in hypothesis generation tasks. *Accounting, Organizations and Society, 20(5),* 345-357

Issa,H., H.Brown-Liburd, and A.Kogan. 2016. *Identifying and Prioritizing Control Deviations Using A Model Derived from Experts' Knowledge* (Working Paper). Rutgers, NJ: Rutgers University.

Jans, M., Alles, M. and 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),* pp.1-20.

Jiambalvo, J., & Waller, W. (1984). Decomposition and assessments of audit risk. *Auditing-A Journal of Practice & Theory, 3(2), 80-88.*

Johnson, J.R., Leitch, R.A. and Neter, J., 1981. Characteristics of errors in accounts receivable and inventory audits. *Accounting Review*, pp.270-293.

Johnstone, D. J. (1994). A statistical paradox in auditing. *Abacus, 30(1),* 44-49.

Johnstone, D. J. (1994). Statistically incoherent hypothesis tests in auditing. *Auditing: A Journal of Practice& Theory*, *14*(2), 156

Joyce, E. J. (1976). Expert judgment in audit program planning. *Journal of Accounting Research*, 29-60

Joyce, E. J., & Biddle, G. C. (1981). Are auditors' judgments sufficiently regressive?. *Journal of Accounting Research*, 323-349.

Joyce, E.J. and Biddle, G.C., 1981. Anchoring and adjustment in probabilistic inference in auditing. *Journal of Accounting Research*, pp.120-145.

Kachelmeier, S.J. and Messier Jr, W.F., 1990. An investigation of the influence of a nonstatistical decision aid on auditor sample size decisions. *Accounting Review*, pp.209-226.

Kaminski, K. A., Sterling Wetzel, T., & Guan, L. (2004). Can financial ratios detect fraudulent financial reporting?. *Managerial Auditing Journal, 19(1),* 15-28

Kaplan, S., Moeckel, C., & Williams, J. D. (1992). Auditors' hypothesis plausibility assessments in an analytical review setting. *Auditing: A Journal of Practice & Theory, 11(2),* 50

Keenoy, C. L. (1958). The impact of automation on the field of accounting. *Accounting Review*, 230-236

Kim, H. S., Neter, J., & Godfrey, J. T. (1987). Behavior of statistical estimators in multilocation audit sampling. *Auditing-A Journal of Practice & Theory, 6(2),* 40-58

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

Kinney Jr, W.R. and Felix Jr, W.L., 1980. Analytical review procedures*. Journal of Accountancy* (pre-1986), 150(000004), p.98.

Kinney Jr, W.R. and Uecker, W.C., 1982. Mitigating the consequences of anchoring in auditor judgments. *Accounting Review*, pp.55-69.

Kinney, W. R. (1975). A decision-theory approach to the sampling problem in auditing. *Journal of Accounting Research,* 117-132.

Kinney, W. R. (1975). Decision theory aspects of internal control system design/compliance and substantive tests. *Journal of Accounting Research*, 14-29.

Kinney, W. R. (1979). Integrating audit tests: Regression analysis and partitioned dollar-unit sampling. *Journal of Accounting Research*, 456-475

Kinney, W. R. (1979). The predictive power of limited information in preliminary analytical review: An empirical study. *Journal of Accounting Research,* 148-165

Kinney, W. R. (1983). A note on compounding probabilities in auditing. *Auditing: A Journal of Practice & Theory, 2(2),* 13-22.

Kinney, W. R. (1986). Audit technology and preferences for auditing standards. *Journal of Accounting and Economics, 8(1),* 73-89.

Kinney, W. R. (1987). Attention-Directing Analytical Review Using Accounting Ratios-A Case-Study. *Auditing-A Journal of Practice & Theory, 6(2),* 59-73.

Kinney, W. R., & Salamon, G. L. (1982). Regression analysis in auditing: A comparison of alternative investigation rules. *Journal of Accounting Research*, 350-366

Knechel, R. (1988). The effectiveness of statistical analytical review on overall audit effectiveness: A simulation analysis. *The Accounting Review*, 74-95

Knechel, W.R., 1985. A stochastic model of error generation in accounting systems. *Accounting and Business Research, 15(59)*, pp.211-221.

Knechel, W.R., 1986. Applications and Implementation A Simulation Study Of The Relative Effectiveness Of Alternative Analytical Review Procedures*. *Decision Sciences, 17(3),* pp.376-394.

Kobelius, J. 2010. The Forrester Wave: Predictive Analytics and Data Mining Solutions. Forrester Research, Inc. USA.

Kochetova-Kozloski, N., & Messier Jr, W. F. (2011). Strategic analysis and auditor risk judgments. *Auditing: A Journal of Practice & Theory, 30(4),* 149-171.

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., 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. and Vasarhelyi, M.A., 1999. Continuous online auditing: A program of research. *Journal of Information Systems, 13(2),* pp.87-103.

Kohavi, R., L. Mason, R. Parekh, and Z. Zheng. 2004. Lessons and Challenges from Mining E-Commerce Data. *Machine Learning,* 57 (1-2): 83-113.

Koonce, L., 199*2. Explanation and counter explanation during audit analytical review.* Accounting Review, pp.59-76.

Kosba, A., Miller, A., Shi, E., Wen, Z., & Papamanthou, C. (2015). *Hawk: The blockchain model of cryptography and privacy-preserving smart contracts*. Cryptology ePrint Archive, Report 2015/675, 2015. http://eprint. iacr. org.

Koskivaara, E. (2004). Artificial neural networks in analytical review procedures. *Managerial Auditing Journal, 19(2),* 191-223.

Koskivaara, E., 2000. Artificial neural network models for predicting patterns in auditing monthly balances. *Journal of the Operational Research Society*, pp.1060-1069.

Krahel J.P. and W.R. Titera. 2015. Consequences of Big Data and Formalization on Accounting and Auditing Standards. *Accounting Horizons*, 29 (2): pp 409-422.

Kreutzfeldt, R.W. and Wallace, W.A., 1986. Error Characteristics In Audit Populations-Their Profile And Relationship To Environmental-Factors. *Auditing: A Journal of Practice & Theory, 6(1),* pp.20-43.

Krishnamoorthy, G., Mock, T. J., & Washington, M. T. (1999). A comparative evaluation of belief revision models in auditing. *Auditing: A Journal of Practice & Theory, 18(2),* 105-127.

Kuenkaikaew, S., & Vasarhelyi, M. A. (2013). The Predictive Audit Framework. *International Journal of Digital Accounting Research*. Vol. 13: 37-71.

Kuhn, J. R., and S. G. Sutton. 2010. Continuous Auditing in ERP System Environments: The Current State and Future Directions. *Journal of Information Systems,* Vol. 24, No. 1: pp. 91-112.

Laney, D. 2001. "3D data management: Controlling data volume, velocity and variety." *META Group Research Note* 6.

Lee, M., Cho, M., Gim, J., Jeong, D. H., & Jung, H. (2014). Prescriptive Analytics System for Scholar Research Performance Enhancement. In *HCI International 2014-Posters' Extended Abstracts* (pp. 186-190). Springer International Publishing.

Lee,M., M.Cho, J.Gim, D.H.Jeong, and Jung, H. 2014. Prescriptive Analytics System for Scholar Research Performance Enhancement. In *HCI International 2014-Posters' Extended Abstracts* (pp. 186-190). Springer International Publishing.

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

Levitan, A. S., & Knoblett, J. A. (1985). Indicators of exceptions to the going concern assumption. *Auditing-A Journal of Practice & Theory*, 5(1), 26-39.

Li, Y., Roge, J.N., Rydl, L. and Hughes, J., 2007. Achieving Sarbanes-Oxley compliance with XBRL-based ERP and continuous auditing. *Issues in Information Systems*, *8*(2), pp.430-436.

Li,H., J.Dai, T.Gershberg, and M.A.Vasarhelyi. 2015. Understanding Usage and Value of Audit Analytics in the Internal Audit: An Organizational Approach. Working paper, Continuous Auditing and Reporting Laboratory, 2013.

Liao, C. and Squicciarini, A., 2015, May. Towards Provenance-Based Anomaly Detection in MapReduce. In *Cluster, Cloud and Grid Computing (CCGrid), 2015 15th IEEE/ACM International Symposium on* (pp. 647-656). IEEE.

Libby, R. (1985). Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research,* 648-667.

Libby, R. and Frederick, D.M., 1990. Experience and the ability to explain audit findings. *Journal of Accounting Research,* pp.348-367.

Libby, R., & Lewis, B. L. (1982). Human information processing research in accounting: The state of the art in 1982. *Accounting, Organizations and Society, 7(3),* 231-285.

Liddy,J. P. 2015. The Future of Audit. Available at: www.forbes.com/sites/realspin/2014/08/04/the-future-of-audit/

Lin, J. and Ryaboy, D., 2013. Scaling big data mining infrastructure: the twitter experience. *ACM SIGKDD Explorations Newsletter*, *14*(2), pp.6-19.

Lin, J. W., Hwang, M. I., & Becker, J. D. (2003). A fuzzy neural network for assessing the risk of fraudulent financial reporting. *Managerial Auditing Journal, 18(8),* 657-665.

Lin, W. T., Mock, T. J., & Wright, A. (1984). The use of the analytic hierarchy process as an aid in planning the nature and extent of audit procedures. *Auditing-A Journal of Practice & Theory, 4(1),* 89-99

Liu, Q .(2014) The application of exploratory Data Analysis in Auditing. *PhD Dissertation*, Rutgers Business School, Continuous Audit and reporting Lab, Newark, NJ, 2014

Loebbecke, J. K. (1995). On the use of Bayesian statistics in the audit process. *Auditing: A Journal of Practice & Theory, 14(2)*, 188

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

Loebbecke, J.K., Eining, M.M. and Willingham, J.J., 1989. Auditors experience with material irregularities-frequency, nature, and detectability. *Auditing: A Journal of Practice & Theory, 9(1)*, pp.1-28.

Lorek, K. S., Branson, B. C., & Icerman, R. C. (1992). On the use of time-series models as analytical procedures. *Auditing: A Journal of Practice & Theory, 11(2),* 66

Lowensohn, S., Johnson, L. E., Elder, R. J., & Davies, S. P. (2007). Auditor specialization, perceived audit quality, and audit fees in the local government audit market. *Journal of Accounting and Public Policy, 26(6),* 705-732

Margo, D., and R. Smogor. 2010. "Using provenance to extract semantic file attributes." In *USENIX 2nd Conf on Theory and practice of provenance*, vol. 7.

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

Martin, R. D., Rich, J. S., & Wilks, T. J. (2006). Auditing fair value measurements: A synthesis of relevant research. *Accounting Horizons, 20(3),* 287-303

McDaniel, L. S., & Simmons, L. E. (2007). Auditors' assessment and incorporation of expectation precision in evidential analytical procedures. *Auditing: A Journal of Practice & Theory, 26(1),* 1-18

McDaniel, P., K. Butler, S. McLaughlin, R. Sion, E. Zadok, and M. Winslett. 2010. "Towards a secure and efficient system for end-to-end provenance." In *Proc. USENIX Workshop on the Theory and Practice of Provenance (TaPP).*

Meservy, R. D., Bailey, A. D., & Johnson, P. E. (1986). Internal control evaluation: A computational model of the review process. *Division of Research, College of Administrative Science, Ohio State University*

Messier Jr, W. F., Kachelmeier, S. J., & Jensen, K. L. (2001). An experimental assessment of recent professional developments in nonstatistical audit sampling guidance. *Auditing: A Journal of Practice & Theory, 20(1),* 81-96

Messier Jr, W. F., Simon, C. A., & Smith, J. L. (2012). Two decades of behavioral research on analytical procedures: What have we learned?. *Auditing: A Journal of Practice & Theory, 32(1),* 139-181

Mittal, A., 2013. Trustworthiness of big data. *International Journal of Computer Applications*, *80*(9).

Mock, T. J., & Willingham, J. J. (1983). An improved method of documenting and evaluating a system of internal accounting controls. *Auditing: A Journal of Practice and Theory, 2(2),* 91-99

Mock, T. J., & Wright, A. M. (1982). Evaluating the effectiveness of audit procedures. *Auditing: A Journal of Practice & Theory, 2(1*), 33-44

Mock, T.J. and Wright, A., (1993). An exploratory study of auditors' evidential planning judgments. *Auditing: A Journal of Practice & Theory, 12(2),* p.39.

Mock, T.J., Wright, A. and Srivastava, R.P., 1998. Audit program planning using a belief function framework. In *Proceedings of the 1998 Deloitte & Touche University of Kansas Symposium on Auditing Problems* (pp. 115-142).

Mogilner, C., T. Rudnick, and S. S. Iyengar. (2008). The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. *Journal of Consumer Research,* 35 (2): 202–15.

Mogis, R. C., & Rogoff, D. (1962). Statistics Offers a Solution to Tomorrow's Auditing Complexities. *Accounting Review*, 704-707.

Montgomery, R.H. 1919. Auditing Theory and Practice, The Ronald Press, 1919 2nd edition, New York

Moreau, L., B. Ludäscher, I. Altintas, R. S. Barga, S. Bowers, S. Callahan, G. Chin et al. (2008b). "Special issue: The first provenance challenge." *Concurrency and computation: practice and experience* 20, no. 5: 409-418.

Moreau, L., J. Freire, J. Futrelle, R. E. McGrath, J. Myers, and P. Paulson. (2008c)."The open provenance model: An overview." In *Provenance and Annotation of Data and Processes*, pp. 323-326. Springer Berlin Heidelberg.

Moreau, L., P. Groth, S. Miles, J. Vazquez-Salceda, J. Ibbotson, S. Jiang, S. Munroe et al. (2008a). "The provenance of electronic data." *Communications of the ACM* 51, no. 4: 52-58.

Muniswamy-Reddy, K .K., P. Macko, and M. I. Seltzer. 2010. "Provenance for the Cloud." In *FAST*, vol. 10, pp. 15-14.

Muniswamy-Reddy, K. K., D. A. Holland, U. Braun, and M. I. Seltzer. 2006. "Provenance-Aware Storage Systems." In *USENIX Annual Technical Conference, General Track*, pp. 43-56.

Muniswamy-Reddy, K. K., U. Braun, D. A. Holland, P. Macko, D. Maclean, D. Margo, M. Seltzer, and R. Smogor. 2009. "Layering in provenance systems." In *Proceedings of the 2009 USENIX Annual Technical Conference*.

Mutchler, J. F. (1984). Auditors' perceptions of the going-concern opinion decision. *Auditing: A Journal of Practice & Theory, 3(2),* 17-30.

Mutchler, J. F. (1985). A multivariate analysis of the auditor's going-concern opinion decision. Journal of Accounting Research, 668-682.

Mutchler, J. F., & Williams, D. D. (1990). The relationship between audit technology, client risk profiles, and the going-concern opinion decision. *Auditing: A Journal Of Practice & Theory, 9(3),* 39-54.

Nearon, B.H., 2005. Foundations in auditing and digital evidence. *The CPA Journal*, *75*(1): p.32.

Nelson, M. K. (1995). Strategies of auditors: Evaluation of sample results. *Auditing: A Journal of Practice and Theory*, *14*(1), 34

Neter, J. (1949). An investigation of the usefulness of statistical sampling methods in auditing. *Journal of Accountancy (pre-1986)*, *87*(000005), 390.

Neter, J., Leitch, R. A., & Fienberg, S. E. (1978). Dollar unit sampling: Multinomial bounds for total overstatement and understatement errors. *Accounting Review*, 77-93

Nichols, D. R., & Baker, R. C. (1977). Testing the Consistency of Auditors' Prior Distributions and Sampling Results. *Abacus*, *13(2),* 91-105

Nigrini, M. J., & Miller, S. J. (2009). Data diagnostics using second-order tests of Benford's law. *Auditing: A Journal of Practice & Theory, 28(2),* 305-324

Nigrini, M. J., & Mittermaier, L. J. (1997). The use of Benford's law as an aid in analytical procedures. *Auditing: A Journal of Practice and Theory, 16(2),* 52

Nogler, G. E. (1995). The resolution of auditor going concern opinions. *Auditing: A Journal of Practice and Theory, 14(2),* 54

O'Driscoll, A., Daugelaite, J. and Sleator, R.D., 2013. 'Big data', Hadoop and cloud computing in genomics. *Journal of biomedical informatics*, *46*(5), pp.774-781.

O'Donnell, E., & Perkins, J. D. (2011). Assessing risk with analytical procedures: Do systems-thinking tools help auditors focus on diagnostic patterns?. *Auditing: A Journal of Practice & Theory, 30(4),* 273-283.

Okab, R. (2013). The Expert Systems and Their Role in Developing External Auditor's Performance and Improving Audit Service's Quality in Information Technology Environment in Audit's Offices Located in the Hashemite Kingdom of Jordan. *International Journal of Business and Management, 8(17)*, p129.

Olavsrud, T. 2016. 21 data and analytic trends that will dominate 2016. Available at: http://www.cio.com/article/3023838/analytics/21-data-and-analytics-trends-that-will-dominate-2016.html

Omoteso, K. (2012). The application of artificial intelligence in auditing: Looking back to the future. *Expert Systems with Applications, 39(9),* 8490-8495.

Oppliger, R. and Rytz, R., 2003. Digital evidence: Dream and reality. *IEEE security & privacy*, (5), pp.44-48.

Paino, H., Hadi, K.A.A. and Tahir, W.M.M.W., 2014. Financial statement error: client's business risk assessment and auditor's substantive test. *Procedia-Social and Behavioral Sciences, 145*, pp.316-320.

Parasuraman, R., T. B. Sheridan, and C.D. Wilkens. 2000. *A Model for Types and Levels of Human Interaction with Automation.* Available at: http://hci.cs.uwaterloo.ca/faculty/elaw/cs889/reading/automation/sheridan.pdf

Park, H., R. Ikeda, and J. Widom. 2011. "Ramp: A system for capturing and tracing provenance in MapReduce workflows."

Park, S., and Y. Lee. 2013. "Secure Hadoop with Encrypted HDFS." In *Grid and Pervasive Computing*, pp. 134-141. Springer Berlin Heidelberg.

Patil, D.J., 2012. *Data Jujitsu: The Art of Turning Data into Product*. O'Reilly Media, Inc.

Pearson, T. A., & Singleton, T. W. (2008). Fraud and forensic accounting in the digital environment. *Issues in Accounting Education*, *23*(4), 545-559.

Peek, L. E., Neter, J., & Warren, C. (1991). AICPA Nonstatistical Audit Sampling Guidelines-A Simulation. *Auditing-A Journal of Practice & Theory, 10(2)*, 33-48

Perols, J and B. Lougee. 2011. The Relation between earnings management and financial statement fraud. *Advances in Accounting, incorporating Advances in Accounting 27 (2011),* pp 19-53

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

Perols,J., and B. Lougee. 2011. The Relation between earnings management and financial statement fraud. *Advances in Accounting*, incorporating Advances in Accounting 27 (2011): pp 19-53

Polato, I., Ré, R., Goldman, A. and Kon, F., 2014. A comprehensive view of Hadoop research—A systematic literature review. *Journal of Network and Computer Applications*, *46*, pp.1-25.

Ponemon, L. A., & Wendell, J. P. (1995). Judgmental versus random sampling in auditing: An experimental investigation. *Auditing: A Journal of Practice & Theory, 14(2),* 17.

Public Company Accounting Oversight Board (PCAOB). 2010. *Audit Evidence.* Auditing Standard (AS) No. 1105. Washington, D.C: PCAOB.

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

Public Company Accounting Oversight Board (PCAOB). 2010. *Audit Evidence.* PCAOB Auditing Standards (AS) 1105. Washington, D.C.: PCAOB

Public Company Accounting Oversight Board (PCAOB). 2010. *Audit Sampling.* Auditing Standard (AS) No. 2315. Washington D.C.: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2010. *Evaluating Audit Results.* Auditing Standard (AS) No. 2810. Washington, D.C.: PCAOB.

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

Public Company Accounting Oversight Board (PCAOB). 2010. *Substantive Analytical Procedures.* Auditing Standard (AS) No. 2305. Washington D.C.: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2016. *Audit Sampling.* Auditing Standards (AS) 2315. Washington, D.C.: PCAOB

Public Company Accounting Oversight Board (PCAOB). 2016. *Substantive Audit Procedures*. Auditing Standards (AS) 2305. Washington, D.C.: PCAOB

Ramage, J. G., Krieger, A. M., & Spero, L. L. (1979). An empirical study of error characteristics in audit populations. *Journal of Accounting Research*, 72-102.

Ratcliffe, T.A. and Munter, P., 2002. Information technology, internal control, and financial statement audits. *The CPA Journal*, *72*(4), p.40

Ravisankar, P., Ravi, V., Rao, G. R., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems*, *50*(2), 491-500

Roberts, D. M. (1974). A statistical interpretation of SAP No. 54. *Journal of Accountancy (pre-1986), 137(000003),* 47

Robertson, J. C. (1995). A constrained cost-minimization model for audit sample planning. *Auditing: A Journal of Practice & Theory, 14(2),* 105

Rohrbach, K. J. (1993). Variance augmentation to achieve nominal coverage probability in sampling from audit populations. *Auditing: A Journal of Theory & Practice, 12(2),* 79

Rohrbach, K. J. (1997). Sample size determination using the augmented variance estimator. *Auditing: A Journal of Practice & Theory, 16(1),* 124

Rosman, A., Biggs, S., Graham, L., & Bible, L. (2007). Successful audit workpaper review strategies in electronic environments. *Journal of Accounting, Auditing & Finance*, *22*(1), 57-83.

Rusitschka, S., Doblander, C., Goebel, C., & Jacobsen, H. A. (2013, December). Adaptive middleware for real-time prescriptive analytics in large scale power systems. In *Proceedings of the Industrial Track of the 13th ACM/IFIP/USENIX International Middleware Conference* (p. 5). ACM.

Sakka, M.A., Defude, B. and Tellez, J., 2010. Document provenance in the cloud: constraints and challenges. In *Networked Services and Applications-Engineering, Control and Management* (pp. 107-117). Springer Berlin Heidelberg.

Scheibehenne, B., R. Greifeneder, and, P. M. Todd. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, 37: 409–425.

Scheidegger, C., D. Koop, E. Santos, H. Vo, S. Callahan, J. Freire, and C. Silva. 2008. "Tackling the provenance challenge one layer at a time." *Concurrency and Computation: Practice and Experience* 20, no. 5: 473-483

Schneider,G., J.Dai, D.Janvrin, K.Ajayi, and R.L. Raschke. 2015. Infer, Predict, and Assure: Accounting Opportunities in Data Analytics. *Accounting Horizons* Vol. 29, No. 3: 719-742.

Scott, W. R. (1973). A Bayesian approach to asset valuation and audit size. *Journal of Accounting Research*, 304-330

Sennetti, J.T., (1990). Toward a more consistent model for audit risk. *Auditing: A Journal Of Practice & Theory, 9(2*), pp.103-112.

Shelton, S. W., Whittington, O. R., & Landsittel, D. (2001). Auditing firms' fraud risk assessment practices. *Accounting Horizons*, *15*(1), 19-33.

Shpilberg, D., Graham, L. E., & Schatz, H. (1986). ExperTAXsm: an expert system for corporate tax planing. *Expert Systems, 3(3),* 136-151

Simmhan, Y. L., B. Plale, and D. Gannon. (2005a). "A survey of data provenance techniques." *Computer Science Department, Indiana University, Bloomington IN* 47405.

Simmhan, Y. L., B. Plale, and D. Gannon. (2005b). "A survey of data provenance in e-science." *ACM Sigmod Record* 34, no. 3: 31-36.

Smieliauskas, W. (1986). A note on a comparison of Bayesian with non-Bayesian dollar-unit sampling bounds for overstatement errors of accounting populations. *Accounting Review*, 118-128

Smith, K. A. (1972). The relationship of internal control evaluation and audit sample size. *Accounting Review*, 260-269

Song, S. K., Jeong, D. H., Kim, J., Hwang, M., Gim, J., & Jung, H. (2014). Research Advising System based on Prescriptive Analytics. In *Future Information Technology* (pp. 569-574). Springer Berlin Heidelberg.

Song, S. K., Kim, D. J., Hwang, M., Kim, J., Jeong, D. H., Lee, S. & Sung, W. (2013, December). Prescriptive Analytics System for Improving Research Power.(2013). In *Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on* (pp. 1144-1145). IEEE.

Song,S. K., D.H. Jeong, J.Kim, M. Hwang, J. Gim, and H. Jung. 2014. Research Advising System based on Prescriptive Analytics. In *Future Information Technology* (pp. 569-574). Springer Berlin Heidelberg.

Song,S. K., D.J. Kim, M. Hwang, J. Kim, D.H. Jeong, S. Lee, and W.Sung. 2013. Prescriptive Analytics System for Improving Research Power. In *Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on* (pp. 1144-1145). IEEE.

Sorensen, J. E. (1969). Bayesian analysis in auditing. *Accounting Review*, 555-561

Souiah, I., A. Francalanza, and V. Sassone. 2009. "A Formal Model of Provenance in Distributed Systems." In *Workshop on the Theory and Practice of Provenance*, pp. 1-11.

Spathis, C., Doumpos, M., & Zopounidis, C. (2003). Using client performance measures to identify pre-engagement factors associated with qualified audit reports in Greece. *The International Journal of Accounting, 38(3),* 267-284

Srinidhi, B. N., & Vasarhelyi, M. A. (1986). Auditor judgment concerning establishment of substantive tests based on internal control reliability. *Auditing: A Journal of Practice & Theory, 5(2).*

Srivastava, R.P., Mock, T.J. and Gao, L., (2011). The Dempster-Shafer Theory: An Introduction and Fraud Risk Assessment Illustration. *Australian Accounting Review, 21(3),* pp.282-291.

Srivastava, R.P., Mock, T.J. and Turner, J.L., (2009). Bayesian fraud risk formula for financial statement audits. *Abacus, 45(1),* pp.66-87.

Srivastava, R.P., Mock, T.J. and Turner, J.L.,(2007). Analytical formulas for risk assessment for a class of problems where risk depends on three interrelated variables. *International Journal of Approximate Reasoning, 45(1),* pp.123-151.

Srivastava, R.P., Wright, A. and Mock, T.J., (2002). Multiple hypothesis evaluation in auditing. *Accounting & Finance, 42(3),* pp.251-277.

Stamatogiannakis, M., Groth, P. and Bos, H., 2014. Looking inside the black-box: Capturing data provenance using dynamic instrumentation. In *Provenance and Annotation of Data and Processes* (pp. 155-167). Springer International Publishing.

Stanley, J. D., & DeZoort, F. T. (2007). Audit firm tenure and financial restatements: An analysis of industry specialization and fee effects. *Journal of Accounting and Public Policy*, *26*(2), 131-159

Stewart, T. (2015). Data analytics for financial-statement Audits, Chapter 5 in AICPA, Audit Analytics and Continuous Audit: Looking Toward the Future, *American Institute Of Certified Public Accountants.* New York, NY 2015

Stewart, T.R. and Kinney Jr, W.R., (2012). Group audits, group-level controls, and component materiality: How much auditing is enough?. *The Accounting Review, 88(2),* pp.707-737.

Strawser, J.R., (1991). Examination of the effect of risk model components on perceived audit risk. *Auditing-A Journal of Practice & Theory, 10(1)*, pp.126-135.

Stringer, K. W. (1975). A statistical technique for analytical review. *Journal of Accounting Research*.

Stringer, K.W., and T.R.Stewart. 1986. *Statistical techniques for analytical review in auditing*. Ronald Press.

Sully, J. M. (1974). Statistical sampling in auditing. *The Statistician*, 71-80.

Tabor, R. H., & Willis, J. T. (1985). Empirical-Evidence On The Changing-Role Of Analytical Review Procedures. *Auditing: A Journal of Practice & Theory, 4(2),* 93-109

Tabor,R. H., and J.T. Willis.1985.Empirical evidence on the changing role of analytical review procedures. *Auditing-a Journal of Practice & Theory*, 4(2), 93-109

Tackett, J., Wolf, F. and Claypool, G., 2004. Sarbanes-Oxley and audit failure: A critical examination. *Managerial Auditing Journal*, *19*(3), pp.340-350.

Tan, W.C. 2004 . "Research Problems in Data Provenance." *IEEE Data Eng. Bull.* 27, no. 4: 45-52.

Tan, W.C. 2007. "Provenance in Databases: Past, Current, and Future." *IEEE Data Eng. Bull.* 30, no. 4 : 3-12.

Taylor, M., J. Haggerty, D. Gresty, and R. Hegarty. 2010. "Digital evidence in cloud computing systems." *Computer Law & Security Review* 26, no. 3: 304-308.

Teeter, R.A. and Vasarhelyi, M.A., 2010. Remote Audit: A Review of Audit-Enhancing Information and Communication Technology Literature. *Journal Of Emerging Technologies In Accounting*.

Teitlebaum, A. D., & Robinson, C. F. (1975). The real risks in audit sampling. *Journal of Accounting Research,* 70-91

Thomas, K., Grier, C., Song, D. and Paxson, V., 2011, November. Suspended accounts in retrospect: an analysis of twitter spam. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference* (pp. 243-258). ACM.

Titera, W.R., 2013. Updating audit standard-Enabling audit data analysis. *Journal Of Information Systems, 27(1),* pp.325-331.

Tracy, J. A. (1969). Bayesian statistical methods in auditing. *Accounting Review*, 90-98

Trompeter, G. and Wright, A. (2010). The World Has Changed - Have Analytical Procedure Practices? *Contemporary Accounting Research* Vol. 27 No.2 (Summer 2010) pp.669-700

Trompeter, G. M., Carpenter, T. D., Desai, N., Jones, K. L., & Riley Jr, R. A. (2012). A synthesis of fraud-related research. *Auditing: A Journal of Practice & Theory, 32(sp1),* 287-321.

Trotman, K., Gramling, A., Johnstone, K., Kaplan, S., Mayhew, B., Reimers, J., Schwartz, R., Tan, H.T., Wright, B., Brazel, J., Earley, C., Krogstad, J., Cohen, J., and Jenkins, G. (2009). Thirty Three Tears of Audit Research. *AAA Audit Section Database*: www2.aaahq.org/audit/33YrsAuditResearchFinal.doc

Tsai, W.T., Wei, X., Chen, Y., Paul, R., Chung, J.Y. and Zhang, D., 2007. Data provenance in SOA: security, reliability, and integrity. *Service Oriented Computing and Applications*, *1*(4), pp.223-247.

Tschakert, N., J. Kokina, S.Kozlowski, and Vasarhelyi, M.A. 2016. CPAs and Data Analytics. The Journal of Accountancy, forthcoming.

Tucker III, J. J. (1994). Initial Efforts of Kenneth W. Stringer to Develop a Statistical Sampling Plan. *The Accounting Historians Journal*, 233-254

Tucker III, J. J., & Lordi, F. C. (1997). Early Efforts of The US Public Accounting Profession to Investigate The Use Of Statistical Sampling. *The Accounting Historians Journal*, 93-116.

Tukey, J. W. 1977. Exploratory data analysis. Reading, MA: Addison-Wesley.

Tukey, J. W. 1980. We need both exploratory and confirmatory. *The American Statistician*, 34: 23-25

Turner, J.L., Mock, T.J. and Srivastava, R.P., (2002). A conceptual framework and case studies on audit planning and evaluation given the potential for fraud. *Business papers*, p.58.

Twitter. 2014. Annual Report 2014. Available at: http://www.viewproxy.com/twitter/2015/1/annualreport2014.pdf

Tysiac, K., 2016. “Internal auditors challenged by cybersecurity, data quality.” Journal of Accountancy, February 16, 2016. Available at: http://www.journalofaccountancy.com/news/2016/feb/internal-audit-challenges-201613894.html

United States Public Law, 2002. Public Law 107-204 107th Congress, *The Sarbanes-Oxley Act of 2002.* Pub. L. No. 107-204, 116 Stat. 745. July 30.

United States Public Law, 2002. Public Law 107-204 107th Congress, *The Sarbanes-Oxley Act of 2002.* Pub. L. No. 107-204, 116 Stat. 745. July 30.

van Aalst, W.M., van Hee, K.M., van Werf, J.M. and Verdonk, M., 2010. Auditing 2.0: using process mining to support tomorrow's auditor. *Computer*, *43*(3), pp.90-93.

Vance, L. L. (1960). A Review of Developments in Statistical Sampling for Accountants. *Accounting Review*, 19-28

Vandervelde, S. D., Chen, Y., & Leitch, R. A. (2008). Auditors' Cross-Sectional and Temporal Analysis of Account Relations in Identifying Financial Statement Misstatements. *Auditing: A Journal of Practice & Theory, 27(2),* 79-107

Vasarhelyi, M. A. (1982). Academic Research in Accounting and Auditing. *Handbook of Accounting and Auditing, hrsg. von John C. Burton, Russel E. Palmer und Robert S. Kay*, *4*.

Vasarhelyi, M. A. (1984). Automation and changes in the audit process. *Auditing: A Journal of Practice & Theory* Vol. 4, No. 1 Fall 1984

Vasarhelyi, M. A. and S. Romero. 2014. Technology in audit engagements: a case study. *Managerial Auditing Journal*, *29*(4), pp.350-365

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

Vasarhelyi, M. A., Bao, D. H., & Berk, J. (1988). Trends in the evolution of scholarly accounting thought: a quantitative examination. *The Accounting Historians Journal*, 45-64.

Vasarhelyi, M. A., Kogan, A., & Tuttle, B. (2015). Big data in accounting: An overview. *Accounting Horizons*.

Vasarhelyi, M.A. 2013. Formalization of Standards, Automation, Robots, and IT Governance. *Journal of Information Systems,* Vol. 27, No. 1 Spring 2013, pp. 1-11.

Vasarhelyi, M.A. 2015. The new scenario of business processes and applications on the digital world. Working Paper, CarLab.

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

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

Vasarhelyi, M.A., and R.Hoitash. 2005. "Intelligent Software Agents in Accounting: An Evolving Scenario." The Evolving Paradigms of Artificial Intelligence and Expert Systems: An International View. Ed. Miklos Vasarhelyi and Alexander Kogan. Vol. 6. Markus Wiener Publishers, 2005

Vasarhelyi, M.A., K.M. Nelson, A.Kogan, R.Srivastava, and H. Lu. 2000. Virtual Auditing Agents: The EDGAR Agent Challenge. Decision Support Systems 28.3. (2000): 241-253.

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

Vaughan, J.A., Jia, L., Mazurak, K. and Zdancewic, S., 2008, June. Evidence-based audit. In *Computer Security Foundations Symposium, 2008. CSF'08. IEEE 21st* (pp. 177-191). IEEE.

Vuchnich, A. (2008). Using CAATTs in preliminary analytical review to enhance the auditor's risk assessment. *CPA JOURNAL*, *78*(5), 38

Wallace, W. A. (1983). The acceptability of regression analysis as evidence in a courtroom–implications for the auditor. *Auditing: A Journal of Practice and Theory, 2(2),* 66-90

Wallace, W.A., (1982). Analytical review developments in practice: misconceptions, potential applications, and field experience. *Working Papers*. Paper 32.

Wang, T., and R. Cuthbertson. 2015. Eight Issues on audit data analytics we would like researched. *Journal of Information Systems* 29 (1): 155-162.

Warren Jr, J.D., Moffitt, K.C. and Byrnes, P., (2015). How Big Data will change accounting. *Accounting Horizons, 29(2),* pp.397-407.

Weber, R. (1978). Auditor decision making on overall system reliability: accuracy, consensus, and the usefulness of a simulation decision aid. *Journal of Accounting Research*, 368-388.

Weitzner, D.J., Abelson, H., Berners-Lee, T., Feigenbaum, J., Hendler, J. and Sussman, G.J., 2008. Information accountability. *Communications of the ACM*, *51*(6), pp.82-87.

Wendell, J. P. (1993). An approach for calculating probabilities for tests of controls using an electronic spreadsheet. *Auditing: A Journal of Practice & Theory, 12(2),* 116

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

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

Wild, J.J. and Biggs, S.F., (1990). Strategic considerations for unaudited account values in analytical review. *Accounting Review,* pp.227-241.

Willingham, J.J. and Wright, W.F., (1985). Financial statement errors and internal control judgments. *Auditing-A Journal of Practice & Theory, 5(1),* pp.57-70.

Wilson, A. C. (1992). The effect of autocorrelation on regression-based model efficiency and effectiveness in analytical review. *Auditing: A Journal of Practice & Theory, 11(1),* 32

Wilson, A. C., & Colbert, J. (1989). An analysis of simple and rigorous decision models as analytical procedures. *Accounting Horizons*, *3*(4), 79.

Wilson, A. C., & Glezen, G. W. (1989). Regression-Analysis In Auditing-A Comparison Of Alternative Investigation Rules-Some Further Evidence. *Auditing: A Journal of Practice & Theory, 8(2)*, 90-100

Wilson, A. C., & Hudson, D. (1989). An empirical study of regression analysis as an analytical procedure*. *Contemporary Accounting Research*, *6*(1), 196-215

Winograd, B. N., Gerson, J. S., & Berlin, B. L. (2000). Audit practices of PricewaterhouseCoopers. *Auditing: A Journal of Practice & Theory, 19(2),* 176-182

Wooldridge, M., and N.R. Jennings.1995. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, *10*(2): 115-152.

Wright, A. and Mock, T.J., (1985). Towards a contingency view of audit evidence. *Auditing: A Journal Of Practice & Theory, 5(1),* pp.91-100.

Wright, A. M., & Bedard, J. C. (2000). Decision processes in audit evidential planning: A multistage investigation. *Auditing: A Journal of Practice & Theory, 19(1),* 123-143.

Wright, A., & Ashton, R. H. (1989). Identifying audit adjustments with attention-directing procedures. *Accounting Review*, 710-728.

Wright, D. W. (1991). Augmenting a sample selected with probabilities proportional to size. *Auditing: A Journal of Practice & Theory, 10(1),* 145-158

Wright, W. F., & Berger, L. (2011). Fraudulent management explanations and the impact of alternative presentations of client business evidence. *Auditing: A Journal of Practice & Theory, 30(2),* 153-171.

Wright, W. F., & Willingham, J. J. (1997). A computational model of loan loss judgments. *Auditing: A Journal of Practice & Theory, 16(1)*, 99

Yoon, K. (2016). “Big Data as Audit Evidence: Utilizing Weather Indicators.” Chapter 3 of the dissertation titled *Three Essays on Unorthodox Audit Evidence,* Rutgers University, Newark N.J.

Yoon, K. 2016. Auditing Revenue Account: Can Audit Sampling Replace Substantive Analytical Procedures? Chapter 2, Dissertation: Three Essays On Unorthodox Audit Evidence. *Rutgers Business School,* Newark NJ

.

Yu, S., & Neter, J. (1973). A stochastic model of the internal control system. *Journal of Accounting Research*, 273-295

Yue,D., X.Wu, Y.Wang, Y.Li, and C.H.Chu. 2007. A review of data mining-based financial fraud detection research. In Wireless Communications, Networking and Mobile Computing, 2007. WiCom. International Conference on (pp. 5519-5522). IEEE.

Zhang, J., Yang, X. and Appelbaum, D., 2015. Toward effective Big Data analysis in continuous auditing. *Accounting Horizons*, *29*(2), pp.469-476.

Zhang,L., A.R.Pawlicki, D.McQuilken, and W.R.Titera. 2012. The AICPA assurance services executive committee emerging assurance technologies task force: The audit data standards (ADS) initiative. *Journal of Information Systems*, *26*(1), 199-205.

Zhao, N., Yen, D.C. and Chang, I.C., 2004. Auditing in the e-commerce era. *Information Management & Computer Security, 12(5),* pp.389-400.

Zikopoulos, P. and Eaton, C., 2011. *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media.

APPENDICES:

Appendix A – Tables and Figures:

Orientation	Techniques:					
D,PD,PS		**Technique Type:**	**E or C**	**S, SS, U**	**QN, QL**	**D, S**
D	**Audit examinations**	Transaction Tests, Ratio Analysis	C	S	QN	D
D		Sampling	C	S	QN	S
D		Confirmations/Re-performance	C	S	QN	D
D		CAATS (IDEA, ACL)	C	S	QN	D
D	**Unsupervised**	Clustering Models	E	S	QN	S
D		Text Mining Models	E	SS,U	QL	S
D		Visualizations	E	SS, U	QL,QN	S
D		Process Mining: Process discovery models	E	S,SS	QN	S
PD	**Supervised**	Process Mining: Process Optimizations	C	S,SS	QN	S

PD		Support Vector Machines (SVM)	C	S	QN	S
PD, PS		Artificial Neural Networks (ANN)	C	S	QN	S
PD, PS		Genetic Algorithms	C	S	QN	S
PD, PS		Expert Systems/Decision Aids	C	S,SS,U	QN,QL	S
PD		Bagging, Boosting models	C	S	QN	S
PD		C4.5 statistical classifiers	C	S	QN	S
PD		Bayesian Theory/Bayesian Belief Networks (BBN)	C	S	QN	S
PD		Dempster-Shafer Theory Models	C	S	QN	S
PD		Probability Theory Models	C	S	QN	S
PD. PS	**Regression**	Log Regression	C	S	QN	S
PD, PS		Linear Regression	C	S	QN	S
PD, PS		Time Series Regression	C	S	QN	S
PD, PS		Auto Regressive Integrated Moving Average (ARIMA)	C	S	QN	S
PD, PS		Univariate and Multivariate Regression Analysis	C	S	QN	S
PD	**Other Statistics**	Multi-criteria Decision Aid	C	S	QN	S
PD		Benford's Law	C	S	QN	S

D		Descriptive Statistics	E	S	QN	S
PD		Structural Models	C	S	QN	S
PD		Analytical Hierarchy Processes (AHP)	C	S	QN	S
D		Spearman Rank Correlation Measurements	E	S	QN	S
PD		Hypothesis Evaluations	C	S	QN	S
PD,PS		Monte Carlo Study/Simulation	C	S	QN	S

Table 20: The Orientation and Techniques of business analytics where:

D, PD,PS = Descriptive, Predictive, Prescriptive; E, C = Exploratory, Confirmatory; S, SS, U = Structured, Semi-Structured, Unstructured; QN, QL = Quantitative, Qualitative; and D,S =Deterministic, Statistical
(adapted from Appelbaum et al 2016)

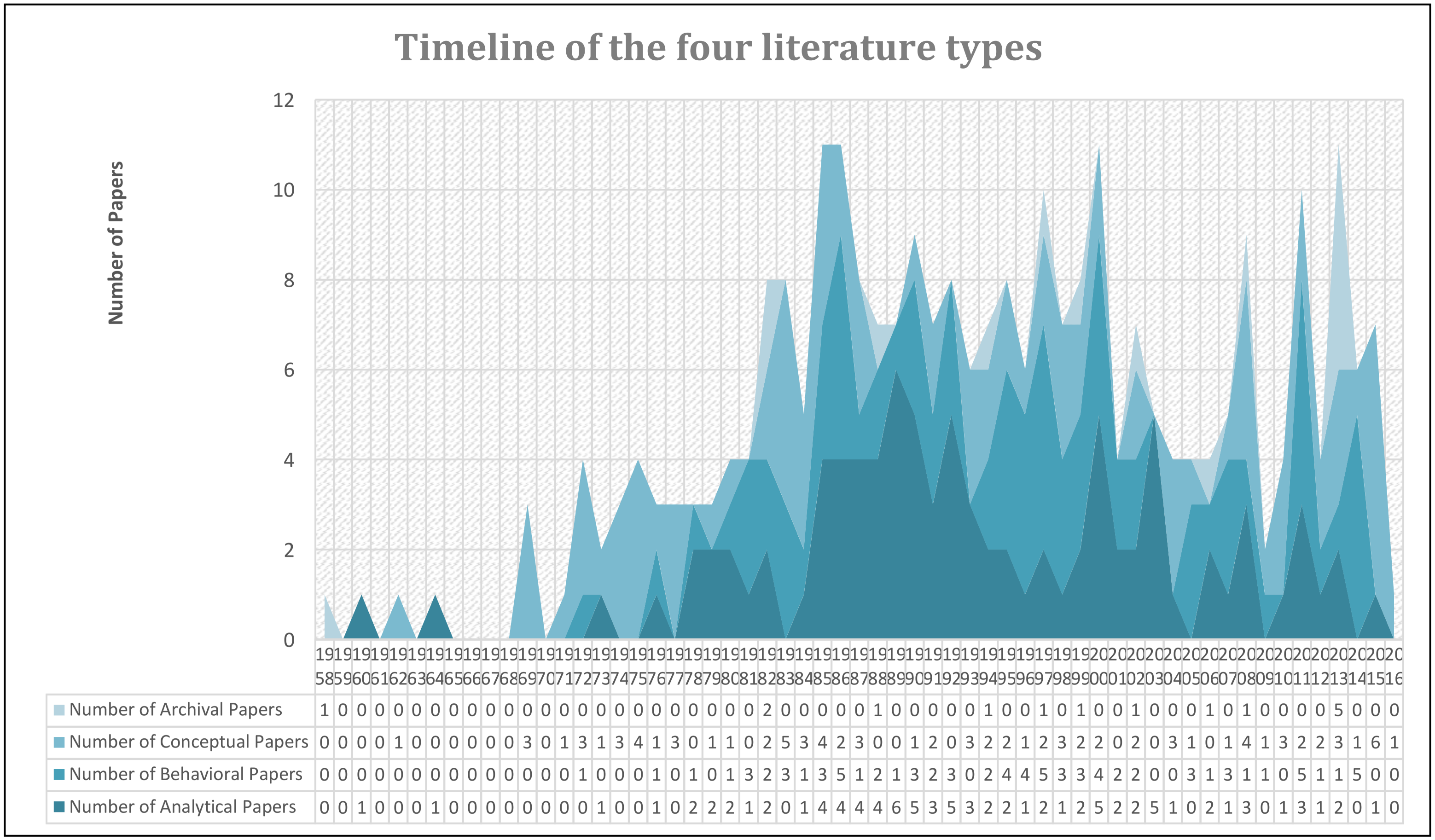

Figure 27: Comparison of the four most heavily used research paper types over time

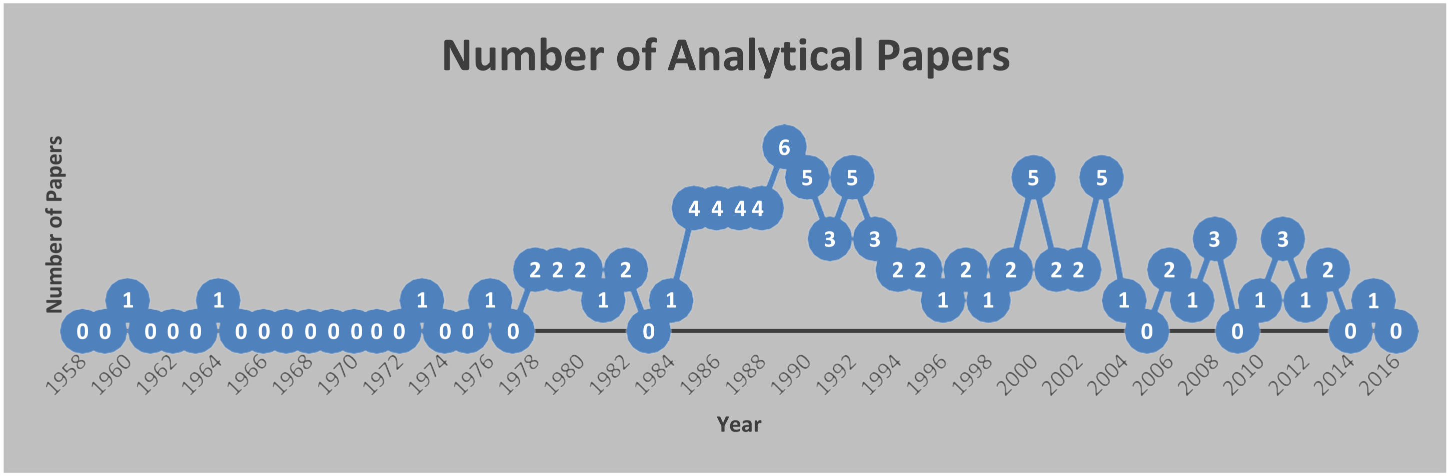

Figure 28: Number of Analytical Papers per year

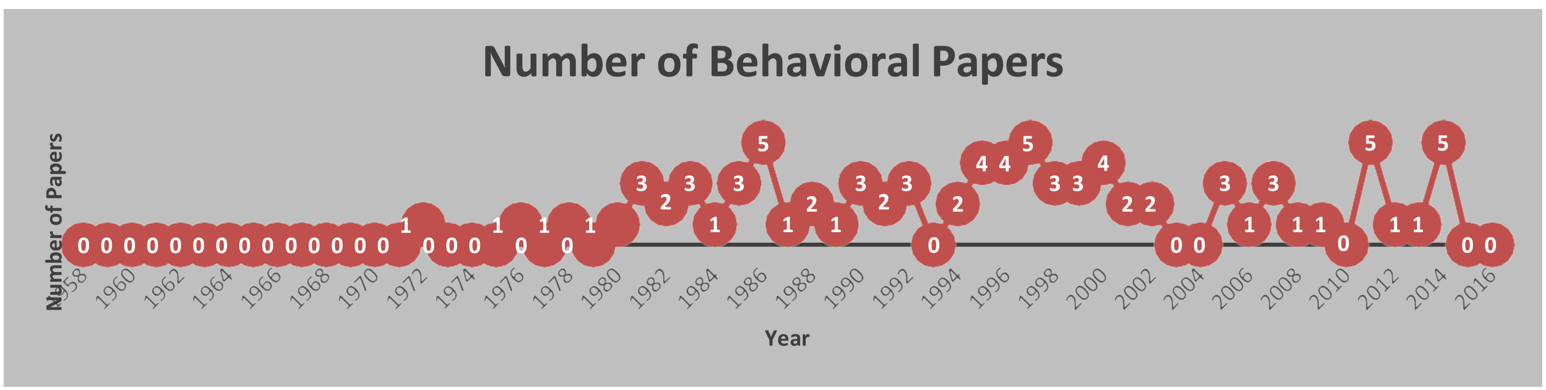

Figure 29: Number of Behavioral Papers that discuss audit analytics in the external audit context

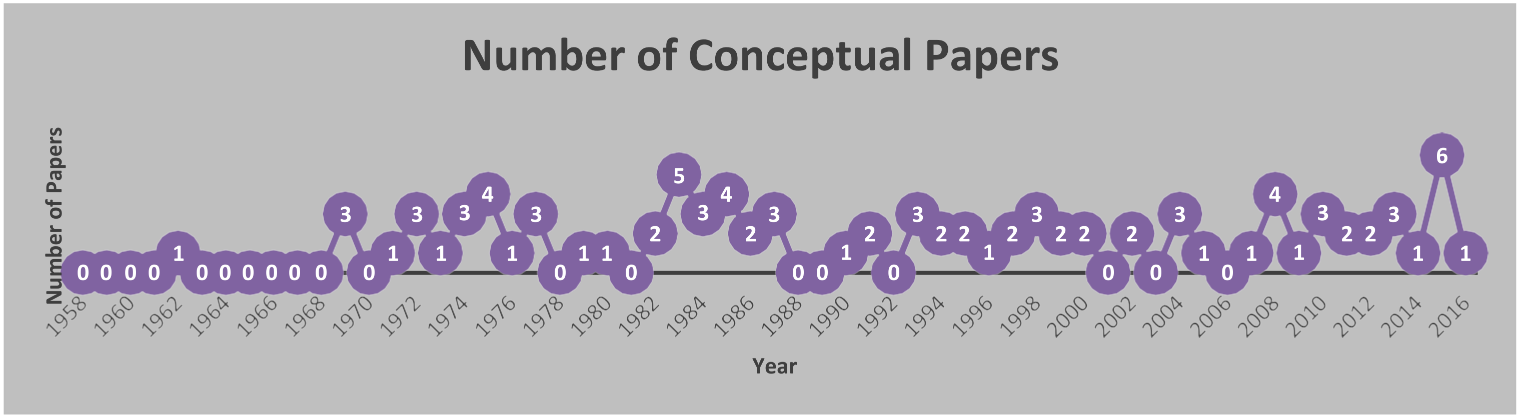

Figure 30: Number of Conceptual Papers that discuss audit analytics in the external audit context

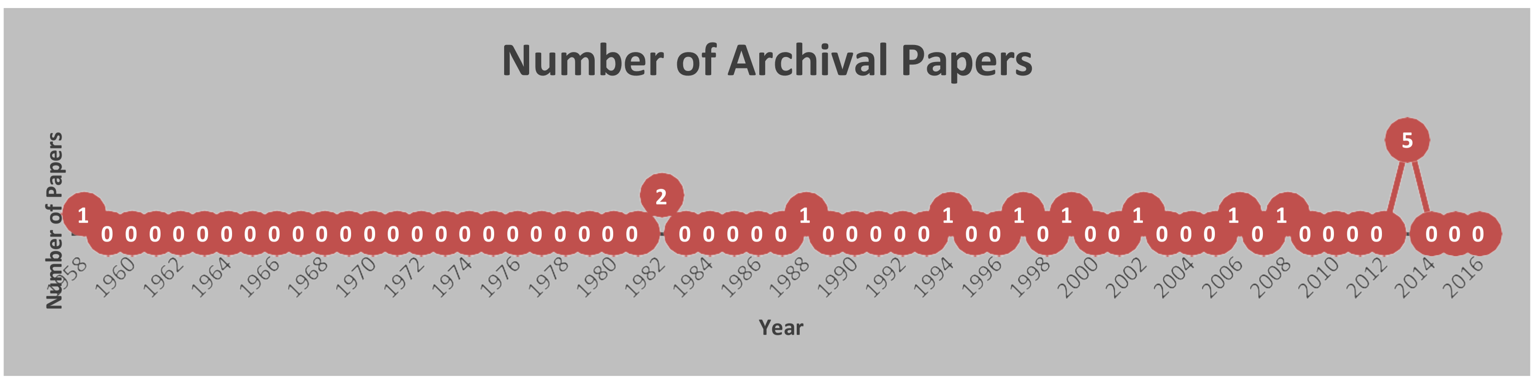

Figure 31: Number of Archival Papers that discuss audit analytics in the external audit context

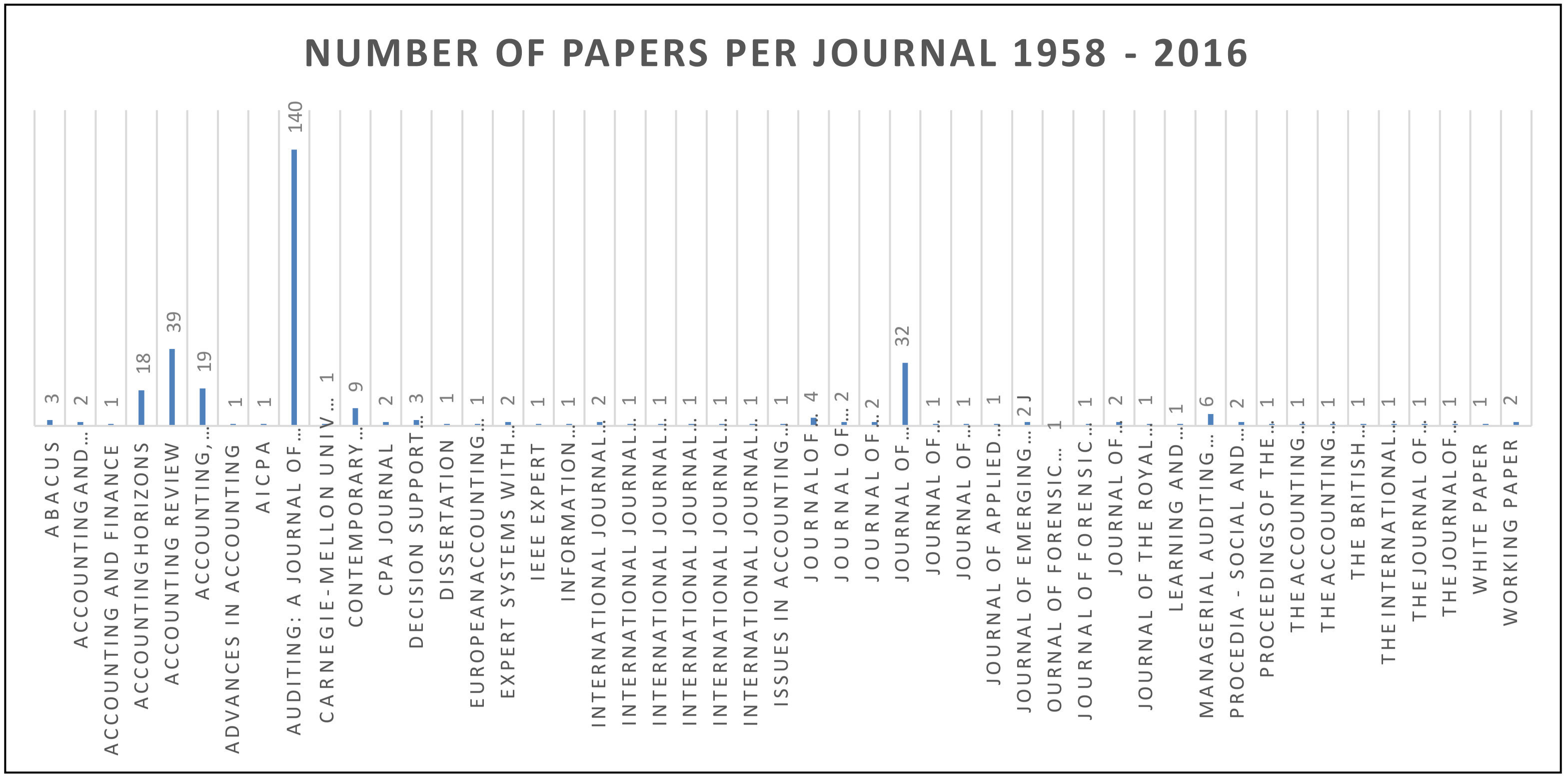

Figure 32: The number of papers published by each journal that discuss Audit Analytics

Techniques:	Audit Examination	Unsupervised	Supervised	Regression	Other Statistics
Audit Phase:					
Engagement:	Ratio Analysis: 17,18,19,22,23,27,29,34,36,39,44,48,49,56,212,213,228,230,237,268, 278	Visualizations: 3, 213, 237	Expert Systems/ Decision Aids: 27, 44, 210, 212, 230, 237, 239	Log Regression: 3, 5, 6, 11, 14, 17, 18, 23, 27, 29, 34, 39, 44, 56, 237	Multi-criteria Decision Aid: 6, 29, 122
		Text Mining: 4, 213, 228, 237		Linear Regression: 12, 14, 39, 49, 56, 237, 278	Structural Models: 56
		Process Mining: 237		Time Series: 34, 56	Descriptive Statistics: 1, 2, 3, 5, 6, 8, 12, 14, 22, 23, 36
				Univariate and Multivariate: 3, 5, 6, 17, 27, 237	
Planning:	Transaction Tests: 15, 16, 17, 19, 20, 26, 44, 116, 126, 177, 213, 215, 230, 235, 236, 237, 238, 254, 263, 274	Clustering: 116, 227, 228, 231, 236, 237	Process Optimization: 58, 227, 236, 237	Log Regression: 1, 2, 3, 5, 7, 10, 11, 14, 16, 17, 18, 20, 21, 23, 27, 29, 32, 33, 34, 37, 38, 39, 41, 49, 50, 55, 56, 68, 74, 83, 100, 106, 112, 113, 114, 116, 118, 123, 124, 142, 143, 144, 145, 146, 147, 149, 150, 151, 153, 154, 156, 168, 173, 215, 221, 223, 227, 231, 233, 234, 237, 270, 292, 300, 301	Multi-criteria Decision Aid: 16, 17, 29, 43, 52, 53, 68, 86, 93, 112, 116, 117, 124, 131, 173
Techniques:	Audit Examination	Unsupervised	Supervised	Regression	Other Statistics

	Ratio Analysis: 16,17,18,19,20,21,22,23,27,29,32,33,34,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53, 58, 61, 62, 66, 67, 68, 71, 75, 76, 79, 80, 86, 87, 93, 95, 97, 100, 102, 103, 106, 109, 110, 113, 114, 117, 121, 124, 125, 128, 130, 131, 132, 133, 134, 135, 136, 137, 139, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 160, 170, 171, 173, 176, 177, 178, 183, 184, 185, 190, 192, 193, 194, 196, 197, 200, 203, 205, 207, 209, 210, 211, 212, 213, 215, 221, 223, 224, 227, 228, 230, 231, 233, 235, 236, 237, 238, 240, 244, 245, 246, 247, 251, 253, 254, 258, 259, 260, 261, 262, 263, 264, 265, 268, 269, 271, 274, 275, 276, 278, 279, 280, 281, 287, 288, 289, 291, 292, 293, 294, 295, 296, 299, 301	Text Mining: 4, 116, 213, 227, 228, 237	Expert Systems/ Decision Aids: 27, 44, 58, 60, 68, 80, 84, 86, 100, 112, 124, 177, 196, 203, 205, 207, 208, 209, 210, 212, 217, 230, 231, 237, 238, 239, 261, 265, 268, 273, 274, 289, 292	Linear Regression: 14, 33, 34, 37, 38, 39, 41, 49, 55, 56, 68, 83, 116, 124, 140, 173, 215, 217, 221, 222, 223, 227, 233, 237, 245, 246, 269, 271, 278, 279, 289, 290, 292, 297, 300, 301	Descriptive Statistics: 1, 2, 3, 5, 7, 8, 9, 10, 12, 13, 14, 22, 23, 32, 36, 37, 116, 124, 143, 144, 145, 147, 155, 173, 227, 244, 246

Techniques:	Audit Examination	Unsupervised	Supervised	Regression	Other Statistics
	CAATS: 20, 26, 28, 44, 63, 116, 126, 137, 139, 155, 177, 213, 215, 230, 237, 238, 254, 265, 274	Visualizations: 3, 10, 116, 213, 227, 228, 237	BBN: 57, 61, 68, 95, 108, 158, 165, 173, 184, 193, 197, 201, 202, 227, 234, 253, 272, 273, 282, 283, 285, 298	Time Series: 30, 33, 34, 38, 41, 55, 56, 66, 68, 83, 87, 105, 116, 120, 124, 140, 168, 173, 183, 215, 217, 222, 223, 227, 237, 238, 265, 269, 271, 287, 289, 292, 300	Structural Models: 56, 105, 118, 140, 269, 287, 300
			Probability Model: 128, 173, 178, 184, 185, 192, 193, 197, 200, 201, 202, 239, 253, 284, 286, 287, 290, 297, 298	ARIMA: 5, 7, 38, 55, 83, 87, 105, 173, 300	
				Univariate and Multivariate: 3, 5, 7, 13, 17, 22, 30, 76, 116, 124, 140, 142, 143, 144, 151, 152, 154, 156, 168, 173, 183, 215, 222, 237, 300	
Techniques:	Audit Examination	Unsupervised	Supervised	Regression	Other Statistics
Substantive & Compliance Testing:	Ratio Analysis: 16,17, 18, 19, 20, 21, 27, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 56, 61, 62, 66, 67, 68, 70, 71, 72, 87, 89, 91, 93, 95, 98, 102, 103, 106, 109, 113, 114, 117, 121, 124,	Visualizations: 3, 116, 133, 177, 213, 229, 237, 238	SVM: 18	Linear Regression: 3, 5, 6, 7, 12, 13, 14, 17, 32, 33, 34, 35, 37, 38, 39, 41, 50, 54, 55, 56, 68, 70, 72, 83, 87, 105, 107, 116, 118, 124, 140, 172, 173, 217, 221, 222, 223, 227, 233, 237, 245, 246, 267, 269, 271, 278, 279, 292, 300, 301	Benford's Law: 28, 116, 119, 124, 126, 136, 238

	125, 130, 131, 132, 133, 134, 136, 139, 141, 142, 143, 144, 145, 146, 147, 149, 150, 151, 153, 154, 155, 160, 170, 171, 173, 174, 175, 176, 177, 178, 183, 184, 185, 188, 190, 193, 194, 196, 197, 200, 211, 212, 213, 216, 221, 223, 227, 228, 229, 230, 233, 235, 236, 237, 238, 240, 241, 242, 243, 244, 245, 246, 247, 248, 251, 252, 254, 258, 259, 260, 262, 264, 265, 266, 267, 269, 271, 274, 276, 278, 279, 280, 281, 287, 292, 294, 301				
Techniques:	Audit Examination	Unsupervised	Supervised	Regression	Other Statistics
	Sampling: 13, 14, 15, 16, 17, 19, 20, 21, 27, 32, 33, 34, 35, 36, 37, 38, 39, 40, 42, 44, 45, 46, 47, 48, 50, 53, 54, 55, 56, 61, 65, 68, 69, 70, 71, 72, 73, 77, 81, 82, 85, 90, 91, 93, 99, 101, 102, 103, 104, 109, 111, 113, 114, 115, 116, 117, 121, 125, 126,	Text Mining: 213, 227, 228, 229, 237	ANN: 17, 18, 27, 211, 227, 233, 237, 279	Time Series: 5, 7, 33, 34, 35, 38, 41, 55, 56, 66, 68, 83, 87, 105, 107, 116, 118, 120, 124, 140, 142, 173, 183, 217, 222, 223, 227, 237, 238, 265, 267, 269, 271, 287, 292, 300	Descriptive Statistics: 3, 5, 6, 7, 8, 12, 13, 14, 32, 36, 37, 116, 124, 143, 144, 145, 147, 155, 173, 227, 229, 241, 244, 246

	127, 130, 131, 132, 133, 136, 138, 142, 143, 144, 145, 146, 147, 149, 150, 151, 153, 154, 157, 160, 161, 162, 163, 164, 165, 166, 167, 169, 170, 173, 174, 175, 176, 177, 178, 183, 184, 189, 193, 194, 195, 197, 198, 199, 200, 202, 203, 213, 216, 220, 221, 222, 223, 225, 226, 227, 228, 229, 230, 233, 237, 238, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 255, 256, 257, 258, 260, 262, 265, 266, 269, 300, 301				
<u>Techniques:</u>	<u>Audit Examination</u>	<u>Unsupervised</u>	<u>Supervised</u>	<u>Regression</u>	<u>Other Statistics</u>
	CAATS: 20, 28, 44, 88, 116, 117, 119, 126, 139, 155, 177, 188, 213, 225, 230, 237, 238, 254, 265, 266, 274	<u>Process Mining: 227, 236, 237, 266</u>	Genetic Algorithms: 18	ARIMA: 5, 7, 35, 38, 55, 87, 105, 107, 118, 173, 267, 300	Structural Models: 56, 105, 118, 140, 269, 287, 300
			Expert Systems/ Decision Aids: 18, 27, 44, 59, 60, 68, 98, 112, 124, 177, 188, 196, 203, 205, 207, 208, 209, 210, 212, 217, 230, 237, 238, 265, 274, 292	Univariate and Multivariate: 3, 5, 6, 7, 13. 17, 35, 90, 116, 124, 140, 142, 143, 144, 151, 154, 156, 173, 183, 222, 137, 300	AHP: 5

			Bagging, Boosting: 77, 173, 227, 260		Monte Carlo Study: 99, 104, 159
			BBN: 57, 61, 68, 89, 90, 95, 108, 111, 156, 158, 159, 162, 163, 164, 165, 173, 176, 184, 185, 193, 195, 197, 198, 199, 201, 202, 227, 234, 298		
			Probability Models: 176, 178, 184, 185, 186, 188, 193, 195, 197, 198, 199, 200, 201, 202, 216, 290, 298		
Techniques:	Audit Examination	Unsupervised	Supervised	Regression	Other Statistics
Review:	Ratio Analysis: 16, 17, 18, 19, 20, 21, 22, 23, 27, 29, 32, 33, 34, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 49, 50, 51, 53, 56, 62, 67, 68, 71, 80, 86, 87, 92, 93, 97, 106, 109, 113, 114, 121, 124, 125, 130, 131, 133, 134, 135, 136, 139, 141, 142, 143, 144, 145, 146, 147, 149, 150, 151, 152, 153, 154, 155, 160, 170, 173, 176, 177, 183, 185, 190, 192, 193, 194, 203, 206, 209, 210, 211, 212, 213, 214, 216, 221,	Visualizations: 3, 10, 116, 133, 177, 213, 237, 238	Expert Systems/ Decision Aids: 18, 23, 27, 44, 59, 60, 68, 80, 86, 112, 124, 177, 203, 208, 209, 210, 212, 217, 230, 232, 237, 238, 239, 289	Linear Regression: 12, 14, 25, 32, 33, 34, 37, 38, 39, 41, 49, 50, 55, 56, 68, 83, 116, 124, 140, 173, 217, 221, 223, 237, 245, 247, 248, 269, 271, 278, 279, 289, 292, 297, 300, 301	Multi-criteria Decision Aid: 6, 16, 17, 29, 53, 68, 86, 93, 112, 116, 124, 131, 173, 143, 144, 145

	223, 224, 228, 230, 232, 236, 237, 238, 240, 241, 243, 245, 246, 247, 262, 264, 269, 271, 276, 278, 281, 287, 289, 292, 293, 299				
Techniques:	Audit Examination	Unsupervised	Supervised	Regression	Other Statistics
	CAATS: 20, 26, 28, 44, 88, 116, 126, 139, 155, 177, 213, 230, 237, 238	Process Mining: 236, 237	BBN: 57, 68, 88, 108	Time Series: 30, 33, 34, 38, 41, 55, 56, 68, 83, 87, 105, 116, 120, 124, 140, 168, 173, 183, 217, 222, 223, 237, 238, 269, 271, 287, 289, 292	Descriptive Statistics: 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 22, 23, 24, 32, 36, 37, 116, 124, 147, 155, 173, 241, 246
			Probability Models: 173, 176, 185, 192, 193, 216, 239, 297, 156, 158, 165, 173, 176, 185, 232, 234	ARIMA: 5, 7, 38, 55, 83, 87, 105, 118, 173, 300	Structural Models: 24, 56, 105, 118, 140, 269, 287
				Univariate and Multivariate: 3, 5, 6, 7, 13, 17, 22, 24, 30, 116, 124, 140, 142, 143, 144, 151, 152, 154, 156, 168, 173, 183, 222, 232, 237, 300	Hypothesis Evaluation: 96
Opinion:	Ratio Analysis: 20,22, 23, 27, 34, 36, 39, 40, 44, 48, 68, 92, 141, 143, 144, 145, 146, 147, 149, 150, 153, 154, 160, 173, 176, 193, 206, 213, 216,	Visualizations: 10, 213, 237	Expert Systems/ Decision Aids:27, 44, 59, 68, 230, 232, 237	Log Regression: 10, 12, 20, 22, 23, 27, 34, 39, 69, 78, 142, 143, 144, 147, 149, 150, 153, 154, 156, 173, 232, 234	Multi-criteria Decision Aid: 68, 78, 173

	228, 230, 232, 237, 243, 278				
		Process Mining: 237		Linear Regression:22, 34, 68, 142, 143, 144, 154, 156, 173, 232, 237	Descriptive Statistics: 9, 10, 12, 23, 36, 143, 144, 145, 147, 147
Techniques:	Audit Examination	Unsupervised	Supervised	Regression	Other Statistics
Continuous Activities:					

Table 21: Each technique per Audit phase is identified with the numbers of all the papers that discuss a particular technique in that phase. The paper numbers are from Table Six in Appendix B

Audit Standard	Potential application areas of EDA
AU-C sec. 240 Consideration of Fraud in a Financial Statement Audit	**.22** Based on analytical procedures performed as part of risk assessment procedures,[8] the auditor should evaluate whether ***unusual or unexpected*** relationships that have been identified indicate risks of material misstatement due to fraud. To the extent not already included, the analytical procedures, and evaluation thereof, should include procedures relating to revenue accounts. (Ref: par. A24–. A26 and .A46) **.27** The auditor should treat those assessed risks of material misstatement due to fraud as significant risks and, accordingly, to the extent not already done so, the auditor should ***obtain an understanding*** of the entity's related controls, including control activities, relevant to such risks, including the evaluation of whether such controls have been suitably designed and implemented to mitigate such fraud risks. (Ref: par.A36–.A37) **.32** Even if specific risks of material misstatement due to fraud are not identified by the auditor, a possibility exists that management override of controls could occur. Accordingly, the auditor should address the risk of management override of controls apart from any conclusions regarding the existence of more specifically identifiable risks by designing and performing audit procedures to, etc. **a.** test the appropriateness of journal entries recorded in the general ledger and other adjustments made in the preparation of the financial statements, including entries posted directly to financial statement drafts. In designing and performing audit procedures for such tests, the auditor should (Ref: par. .A47–.A50 and.A55) i. obtain an ***understanding*** of the entity's financial reporting process and controls over journal entries and other

	adjustments, [12] and the suitability of design and implementation of such controls; ii. make inquiries of individuals involved in the financial reporting process about inappropriate or ***unusual*** activity relating to the processing of journal entries and other adjustments; ,etc.. **c.** evaluate, for significant transactions that are ***outside the normal*** course of business for the entity or that otherwise appear to be ***unusual*** given the auditor's ***understanding*** of the entity and its environment and other information obtained during the audit, whether the business rationale (or the lack thereof) of the transactions suggests that they may have been entered into to engage in fraudulent financial reporting or to conceal misappropriation of assets. (Ref: par. .A54) **.A21** Those charged with governance of an entity oversee the entity's systems for monitoring risk, financial control, and compliance with the law. In some circumstances, governance practices are well developed, and those charged with governance play an active role in oversight of the entity's assessment of the risks of fraud and of the relevant internal control. Because the responsibilities of those charged with governance and management may vary by entity, it is important that the auditor ***understands*** the respective responsibilities of those charged with governance and management to enable the auditor to obtain an understanding of the oversight exercised by the appropriate individuals. **.A37** It is, therefore, important for the auditor to ***obtain an understanding*** of the controls that management has designed, implemented, and maintained to prevent and detect fraud. **.A49** When identifying and selecting journal entries and other adjustments for testing and determining the appropriate method of examining the underlying support for the items selected, the following matters may be relevant: • *The characteristics of fraudulent journal entries or other adjustments*. Inappropriate journal entries or other adjustments often have unique identifying characteristics. Such characteristics may include entries (*a*) made to ***unrelated, unusual, or seldom-used*** accounts***;*** (*b*) made by individuals who typically do not make journal entries; (*c*) recorded at the end of the period or as post closing entries that have little or no explanation or description; (*d*) made either before or during the preparation of the financial statements that do not have account numbers; or (*e*) containing round numbers or consistent ending numbers. • *The nature and complexity of the accounts*. Inappropriate journal entries or adjustments may be applied to accounts that (*a*) contain transactions that are ***complex or unusual*** in nature, (*b*) contain significant estimates and period-end adjustments, (*c*) have been prone to misstatements in the past, (*d*) have not been reconciled on a timely basis or contain unreconciled differences, (*e*) contain intercompany transactions, or (*f*) are otherwise associated with an identified risk of material misstatement due to fraud. In audits of entities that have several locations or components, consideration is given to the need to select journal entries from multiple locations.

Table 22: Potential application areas of EDA in Clarified Statements on Audit Standards issued by the AICPA

Appendix B:

Paper Number	Title	Authors	Year	Journal	Paper Category	Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report	Techniques: Audit Examinations	Techniques: Unsupervised	Techniques: Supervised	Techniques Regression	Techniques: Other Statistics
1	The Relation between earnings management and financial statement fraud	Perols, J.	2011	Advances in Accounting, incorporating Advances in Accounting	analytical	E, P			SVM, ANN, bagging	log regression	descriptive statistics
2	Audit Analytical Procedures: A Field Investigation	D. Eric Hirst and Lisa Koonce	1996	Contemporary Accounting Research	field study	P,ST,REV	sampling, ratio analysis			log regression, linear regression	descriptive statistics
3	Predicting Material Accounting Misstatements	Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G.	2011	Contemporary Accounting Research	analytical	E, P, ST, REV		Visualization		univariate regression analysis, multivariate regression analysis,	descriptive statistics

4	A computational model for financial reporting fraud detection	Glancy, F. H., & Yadav, S. B.	2011	Decision Support Systems	design science	E, P		text mining			
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
5	An application of analytical hierarchy process to model expert judgments on analytical review procedures	Arrington, C. E., Hillison, W., & Jensen, R. E.	1984	Journal of Accounting Research, 298-312.	analytical	E, P, ST, REV				log regression, Box Jenkins, Random Walk, Random Walk Drift, multivariate regression analysis,	descriptive statistics, AHP

6	The Impact of Analytical Review Results, Internal Control Reliability, and Experience on Auditor's Use of Analytical Review	Jeffrey Cohen and Thomas Kida	1989	Journal of Accounting Research	survey	P,ST,REV	sampling, ratio analysis			log regression, linear regression	descriptive statistics
7	ARIMA and regression in analytical review: An analytical test	Kinney Jr, W.R.	1978	Accounting Review	analytical	P,ST,REV				log regression, ARIMA, OLS, Martingale, Sub-Martingale, descriptive statistics, univariate regression analysis	
8	Audit detection of financial statement errors.	Hylas, R. E., & Ashton, R. H.	1982	Accounting Review	behavioral	E,P,ST,REV					descriptive statistics
<u>Paper Number</u>	**<u>Title</u>**	**<u>Authors</u>**	**<u>Year</u>**	**<u>Journal</u>**	**<u>Paper Category</u>**	**<u>Stage(s) of Audit Cycle: E = engagement, P = planning,</u>**	**<u>Techniques: Audit Examinations</u>**	**<u>Techniques: Unsupervised</u>**	**<u>Techniques: Supervised</u>**	**<u>Techniques Regression</u>**	**<u>Techniques: Other Statistics</u>**

						ST = substantive testing, REV = review, R = report					
9	Audit technology and preferences for auditing standards.	Kinney, W. R.	1986	Journal of Accounting and Economics	analytical	P,REV,R					descriptive statistics, Spearman Rank Correlation Measurement
10	Audit effort and earnings management	Caramanis, C., & Lennox, C.	2008	Journal of Accounting and Economics	analytical	P,REV,R		Visualization		log regression	descriptive statistics
11	Analytical formulas for risk assessment for a class of problems where risk depends on three interrelated variables	Rajendra P. Srivastava *, Theodore J. Mock, Jerry L. Turner	2006	InternationalJournal of Approximate Reasonong	conceptual	P	Bayesian Belief Networks, Dempster Shafer Theory				

12	Audit firm tenure and financial restatements: An analysis of industry specialization and fee effects	Stanley, J. D., & DeZoort, F. T.	2007	Journal of Accounting and Public Policy	analytical	E,P,REV				log regression	
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
13	Availability and the generation of hypotheses in analytical review	Libby, R.	1985	Journal of Accounting Research	behavioral	P, ST, REV	sampling			multivariate regression analysis	descriptive statistics
14	Auditors' perceived business risk and audit fees: Analysis and evidence	Bell, T. B., Landsman, W. R., & Shackelford, D. A.	2001	Journal of Accounting research	survey	E, P, ST, REV	sampling			log regression, linear regression, OLS	descriptive statistics

15	The World Fas Changed - Have Analytical Procedure Practices?	Greg Trompeter and Arnold Wright	2010	Contemporary Accounting Research	survey	P, ST, REV	sampling, transaction tests, data analytics				Benfords Law
16	Detecting falsified financial statements: a comparative study using multicriteria analysis and multivariate statistical techniques	Ch. Spathis, M. Doumpos and C. Zopounidis	2002	European Accounting Review	analytical	P, ST, REV	sampling, ratio analysis, transaction tests			log regression	multicriteria decision aid
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**

17	Detecting false financial statements using published data: some evidence from Greece	Ch. Spathis	2002	Managerial Auditing Journal	analytical	E,P,ST,REV	sampling, ratio analysis, transaction tests		ANN	log regression, step-wise logistic, univariate regression analysis, multivariate regression analysis	multicriteria decision aid
18	Detection of financial statement fraud and feature selection using data mining techniques	P. Ravisankar, V. Ravi, G. Rhaghava Rao, and I. Bose	2010	Decision Support Systems	analytical	E,P,ST,REV,	ratio analysis		SVM, ANN, Genetic Algorithm, MLFF, GMDH	log regression	
19	Fraud and Forensic Accounting in the Digital Age	Timothy Pearson and Tommie W. Singleton	2008	Issues in Accounting Education	conceptual	E,P,ST,REV	sampling, ratio analysis, transaction tests				
20	Innovation and Practice of continuous auditing	David Y. Chan and Miklos A. Vasarhelyi	2010	International Journal of Accounting Information Systems	conceptual	P,ST,REV,R	sampling, ratio analysis, transaction tests, data modeling, data analytics			log regression	

Paper Number	Title	Authors	Year	Journal	Paper Category	Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report	Techniques: Audit Examinations	Techniques: Unsupervised	Techniques: Supervised	Techniques Regression	Techniques: Other Statistics
21	Integrating Audit Tests: Regression Analysis and Partitioned Dollar-Unit Sampling	W.R. Kinney	1979	Journal of Accounting Research	conceptual	P,ST,REV,R	sampling, ratio analysis			log regression, OLS	
22	Predicting Audit Qualifications with Financial and Market Variables	Nicholas Dopuch, Robert W. Holthausen and Richard W. Leftwich	1987	Accounting Review	analytical	E,P,ST,REV,R	sampling, ratio analysis			probit model, multivariate regression analysis	descriptive statistics
23	Predicting going concern opinion with data mining	David Martens, Liesbeth Bruynseels, Bart Baesens, Marleen Willekins,	2008	Decision Support Systems	analytical	E,P,REV,R	sampling, ratio analysis		SVM, majority vote, AntMiner+, C4.5 Statistical Classifier	log regression	descriptive statistics

		and Jan Vanthienen									
24	The Prediction Performance of a Structural Model of Accounting Numbers	John J. Wild	1987	Journal of Accounting Research	analytical	REV				step-wise logistic, univariate regression analysis, multivariate regression analysis	descriptive statistics, Structural Model
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

25	Regression Analysis in Auditing: A Comparison of Alternative Investigation Rules	William R. Kinney and Gerald L. Salaman	1982	Journal of Accounting Research	analytical	REV,R				linear regression	
26	The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data	Cindy Durtschi, William Hillison, and Carl Pacini	2004	Journal of Forensic Accounting	conceptual	P,REV,R	transaction tests, data analytics				Benford's Law
27	The application of artificial intelligence in auditing: Looking back to the future	Kamil Omoteso	2012	Expert Systems with Applications	conceptual	E,P,ST,REV,R	sampling, ratio analysis		ANN, expert systems/decision aids	log regression	
28	Evaluation of competing hypotheses in auditing	Stephen K. Asare and Arnold M. Wright	1997	Auditing: A Journal of Practice & Theory	behavioral	P,REV			Process Optimization		Complementary Hypothesis Evaluation, Independent Hypothesis Evaluation

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
29	Using client performance measures to identify pre-engagement factors associated with qualified audit reports in Greece	Charalambos Spathis, Michael Doumpos, and Constantin Zopounidis	2003	The International Journal of Accounting	analytical	E,P,REV	ratio analysis			log regression, discriminant analysis	multicriteria decision aid
30	The role of problem representation shifts in auditor decision processes in analytical procedures.	James L. Bierstaker, Jean C. Bedard, and Stanley F. Biggs	1999	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	ratio analysis			log regression	

31	Evidence on the Effect of Financial and Nonfinancial Trends on Analytical Review	Jeffrey R. Cohen, Ganesh Krishnamoorthy, and Arnold M. Wright	2000	Auditing: A Journal of Practice & Theory	experimental	P,ST,REV	sampling, ratio analysis			log regression, linear regression	
32	The Use of Analytical Procedures	Edward Blocher and George F. Patterson	1996	Journal of Accountancy	conceptual	E,P,ST,REV,R	sampling, ratio analysis				descriptive statistics
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
33	Establishing Investigation Thresholds for Preliminary Analytical Procedures	Robery M. Harper, Jr., Jerry R. Strawser, and Kwei Tang	1990	Auditing: A Journal of Practice & Theory	analytical	P	ratio analysis			multivariate distribution	

34	A Comparison of Analytical Procedure Expectation Models Using Both Aggregate and Disaggregate Data	Simon C. Dzeng	1994	Auditing: A Journal of Practice & Theory	case study	ST	sampling, ratio analysis			log regression, time series regression, ARIMA, Random Walk, multivariate regression analysis	
35	A Reexamination of Auditor versus Model Accuracy within the Context of the Going-Concern Opinion Decision	William Hopwood, James C. McKeown and Jane F. Mutchler	1994	Contemporary Accounting Research	analytical	P,ST,REV,R	sampling, ratio analysis			log regression, univariate regression analysis, multivariate regression analysis	
36	Experience and Auditor's Selection of Relevant Information for Preliminary Control Risk Assessments	Jefferson T. Davis	1996	Auditing: A Journal of Practice & Theory	experimental	P			ANN		

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
37	Sampling Information in Strategic Audit Settings	John C. Fellingham, D. Paul Newman, and Evelyn R. Patterson	1989	Auditing: A Journal of Practice & Theory	analytical	ST	sampling, ratio analysis		Simulation, Process Optimization	linear regression	multicriteria decision aid
38	The Use of Analytical Procedures in Review and Audit Engagements	Frank P. Daroca and William W. Holder	1985	Auditing: A Journal of Practice & Theory	survey	P,ST,REV	sampling, ratio analysis			log regression, linear regression, time series regression, Box Jenkins	

39	A Risk Analysis Approach to Auditing	Michael Grobstein and Paul W. Craig	1984	Auditing: A Journal of Practice & Theory	discussion	E,P,ST,REV,R	sampling, ratio analysis, firm developed proprietary software			log regression, linear regression	
40	Auditor's Perceptions of the Going-Concern Opinion Decision	Jane F. Mutchler	1984	Auditing: A Journal of Practice & Theory	survey	P,ST,REV,R	sampling, ratio analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
41	A Note on the Practice of Analytical Review	Stanely F. Biggs and John J. Wild	1984	Auditing: A Journal of Practice & Theory	survey	P,ST,REV,R	ratio analysis			log regression, linear regression, time series regression	

42	Decomposition and Assessments of Audit Risk	James Jiambalvo and William Waller	1984	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	sampling, ratio analysis				
43	The Use of the Analytic Heirarchy Process as an Aid in Planning the Nature and Extent of Audit Procedures	W. Thomas Lin, Theodore J. Mock and Arnold Wright	1984	Auditing: A Journal of Practice & Theory	theoretical	P	sampling, ratio analysis				multicriteria decision aid
44	Automation and Changes in the Audit Process	Miklos A. Vasarhelyi	1984	Auditing: A Journal of Practice & Theory	theoretical	E,P,ST,REV,R	sampling, ratio analysis, firm developed proprietary software, transaction tests, data modeling		expert systems/decision aids		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST =**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

						substantive testing, REV = review, R = report					
45	Analytical Approaches to Audit risk: A Survey and Analysis	Barry E. Cushing and James K. Loebbecke	1983	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV	sampling, ratio analysis				
46	Auditor's Analytical Review Judgments for Payroll Expense	Edward Blocher, Robert S. Esposito, and John J. Willingham	1983	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	sampling, ratio analysis				
47	Normal Form Decision Theory Development of the Audit Sampling Model	D.R. Finley	1983	Auditing: A Journal of Practice & Theory	conceptual	P,ST,REV	sampling, ratio analysis				
48	Unique Audit Methods: Peat Marwick International	Robert K. Elliott	1983	Auditing: A Journal of Practice & Theory	discussion	E,P,ST,REV,R	sampling, ratio analysis, firm developed proprietary software				

49	A Note on Compounding Probabilities in Auditing	William R. Kinney, Jr.	1983	Auditing: A Journal of Practice & Theory	conceptual	E,P,ST,REV	sampling, ratio analysis			log regression, linear regression	
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
50	The Acceptability of Regression Analysis as Evidence in a Courtroom - Implications for the Auditor	Wanda A. Wallace	1983	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV	sampling, ratio analysis			log regression, linear regression	
51	Analytical Review Procedures in Planning the Audit: An Application Study	William W. Holder	1983	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	ratio analysis				

52	An Improved Method of Documenting and Evaluating A System of Internal Accounting Controls	Theodore J. Mock and John J. Willingham	1983	Auditing: A Journal of Practice & Theory	field study	P	sampling, ratio analysis, firm developed proprietary software				multicriteria decision aid
53	Evaluating the Effectiveness of Audit Procedures	Theodore J. Mock and Arnold Wright	1982	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV	sampling, ratio analysis				multicriteria decision aid
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
54	Ex Post Sampling Risks and Decision Rule Choice in Substantive Testing	Paul J. Beck and Ira Solomon	1985	Auditing: A Journal of Practice & Theory	theoretical	ST	sampling, ratio analysis			log regression, linear regression	multicriteria decision aid

55	The X-11 Model: A New Analytical Review Technique for the Auditor	Michael T. Dugan, James A. Gentry and Keith A. Shriver	1985	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV	sampling			log regression, linear regression, time series regression, ARIMA, Box Jenkins	
56	Empirical Evidence on the Changing Role of Analytical Review Procedures	Richard H. Tabor and James T. Willis	1985	Auditing: A Journal of Practice & Theory	survey	E,P,ST,REV	sampling, ratio analysis			log regression, linear regression, time series regression	Structural Model
57	The Development of Bayesian Decision-Theoretic Concepts in Attribute Sampling	Michael A. Crosby	1985	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV			Bayesian Theory/Bayesian Belief Networks, BSTS, Naïve Bayes		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing,**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**

						<u>REV = review, R = report</u>					
58	Internal Control Evaluation: A Computational Model of the Review Process	Rayman D. Meservy, Andrew D. Bailey, Jr. and Paul E. Johnson	198 6	Auditing: A Journal of Practice & Theory	theoretic al	P			expert systems/decisi on aids, Simulation, Process Optimization		
59	Developing ExperTAX: An Expert System for Corporate Tax and Accrual Planning	David Shpilberg and Lynford E. Graham	198 6	Auditing: A Journal of Practice & Theory	case study	P,ST,REV ,R	firm develop ed propriet ary software		expert systems/decisi on aids		
60	A Preliminary Investigation of EDP-XPERT	James V. Hansen and William F. Messier, Jr.	198 6	Auditing: A Journal of Practice & Theory	behavior al	P,ST,REV	firm develop ed propriet ary software		expert systems/decisi on aids		
61	Efficiency of the Bayesian Approach in Compliance Testing: Some Empirical Evidence	Mohammad J. Abdolmoha mmadi	198 6	Auditing: A Journal of Practice & Theory	behavior al	P,ST	samplin g, ratio analysis		Bayesian Theory/Bayesi an Belief Networks		

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**
62	The Impact of Advanced Systems on Controls and Audit Procedures: A Theory and an Empirical Test	Gordon B. Davis and Ron Weber	1986	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	ratio analysis		Simulation, Process Optimization		
63	Auditor Judgment concerning establishment od substantive tests based on internal control reliability	B.N. Srinidi and M.A. Vasarhelyi	1986	Auditing: A Journal of Practice & Theory	behavioral	P	sampling, data modeling		Simulation, Process Optimization		
64	A continuous constrained optimization model for audit sampling	D.R. Finley and J.L. Boockholdt	1987	Auditing: A Journal of Practice & Theory	theoretical	P			Continuous Constrained Optimization Model		

65	Behavior of Statistical Estimators in Multilocation Audit Sampling	Hyo Seuk Kim, John Neter, and James T. Godfrey	1987	Auditing: A Journal of Practice & Theory	theoretical	ST	sampling				
66	An Investigation of the Use of Preliminary Analytical Review to Provide Substantive Audit Evidence	James K. Loebbecke and Paul J. Steinbart	1987	Auditing: A Journal of Practice & Theory	analytical	P,ST	ratio analysis			time series regression, Sub-Martingale	
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**
67	A Study of Auditors' Analytical Review Performance	Edward Blocher and Jean C. Cooper	1988	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	ratio analysis				

68	Research in Internal Control Evaluation	William L. Felix Jr and Marcia S. Niles	1988	Auditing: A Journal of Practice & Theory	literature review	P,ST,REV,R	sampling, ratio analysis, firm developed proprietary software		expert systems/decision aids, Bayesian Theory/Bayesian Belief Networks, Simulation, Process Optimization	log regression, linear regression, time series regression	multicriteria decision aid
69	Sample Size Planning for the Moment Method of MUS: Incorporating Audit Judgments	Richard A. Grimlund	1988	Auditing: A Journal of Practice & Theory	analytical	ST	sampling		Simulation, Process Optimization		multicriteria decision aid
70	On the Robustness of Model-Based Sampling in Auditing	Chen-en Ko, Christopher J. Nachtsheim, Gordon L. Duke and Andrew D. Bailey, Jr.	1988	Auditing: A Journal of Practice & Theory	analytical	ST	sampling, ratio analysis			linear regression	
71	Analytical Procedures: A Defensive Necessity	Frank Coglitore and R. Glen Berryman	1988	Auditing: A Journal of Practice & Theory	case study	P,ST,REV	sampling, ratio analysis				

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
72	Planning Sample Sizes for Stringer-Method Monetary Unit and Single-Stage Attribute Sampling Plans	Thomas W. Hall, Bethane Jo Pierce, and W.R. Ross	1989	Auditing: A Journal of Practice & Theory	analytical	ST	sampling				
73	Regression Analysis in Auditing: A comparison of Alternative Investigation Rules-some further evidence	Arlette C. Wilson abd G. William Glezen	1989	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV				log regression	
74	An archival investigation of audit program planning	Jean C. Bedard	1989	Auditing: A Journal of Practice & Theory	survey	P	ratio analysis				

75	An analysis of simple and rigorous decision models as analytical procedures	Arlette C. Wilson and Janet Colbert	1989	Accounting Horizons	analytical	P,ST,REV			expert systems/decision aids	linear regression, time series regression	
76	Computer-Intensive Methods in Auditing: Bootstrap Difference and Ratio Estimation	Gary C. Biddle, Carol M. Bruton, and Andrew F. Siegel	1990	Auditing: A Journal of Practice & Theory	analytical	ST	sampling		Bootstrap		
77	The Relationship Between Audit Technology, Client Risk Profiles, and the Going-Concern Opinion Decision	Jane F. Mutchler and David D. Williams	1990	Auditing: A Journal of Practice & Theory	analytical	REP	log regression				multicriteria decision aid
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV =**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

						review, R = report					
78	An Investigation of the Influence of a Nonstatistical Decision Aid on Auditor Sample Size Decisions	Steven J. Kachelmeier and William F. Messier, Jr.	1990	The Accounting Review	behavioral	ST	Sampling, Ratio Analysis				
79	Developing Risk Advisor: An Expert System for Risk Identification	Lynford E. Graham, Jeffrey Damens and George Van Ness	1991	Auditing: A Journal of Practice & Theory	case study	P,REV,R	ratio analysis, firm developed proprietary software		expert systems/decision aids		
80	Augmenting a Sample Selected with probablities Proportinal to Size	David W. Wright	1991	Auditing: A Journal of Practice & Theory	analytical	ST	sampling				
81	AICPA Nonstatistical Audit Sampling Guidelines: A Simulation	Lucia E. Peek, John Neter, and Carl Warren	1991	Auditing: A Journal of Practice & Theory	survey	ST	sampling				

82	analytical analysis of audit uncertainty qualifications.	Timothy B. Bell and Richard H. Tabor	1991	Journal of Accounting Research	analytical	P,ST,REV,R	sampling, ratio analysis			log regression, multivariate regression analysis	descriptive statistics
83	Evaluating Expert Systems with Complex Outputs: The Case of Audit Planning	J. Efrim Boritz and Anthony K.P. Wensley	1992	Auditing: A Journal of Practice & Theory	case study	P			expert systems/decision aids		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
84	Auditor's Hypothesis Plausibility Assessments in an Analytical Review Setting	Steven E. Kaplan, Cindy Moeckel,and Joanne Deahl Williams	1992	Auditing: A Journal of Practice & Theory	behavioral	P,REV	ratio analysis, firm developed proprietary software		expert systems/decision aids		multicriteria decision aid

85	Strategies of Audtors: Evaluation of Sample Results	Martha Krug Nelson	1995	Auditing: A Journal of Practice & Theory	behavioral	ST	sampling, ratio analysis				
86	On the Use of Time-Series Models as Analytical Procedures	Kenneth S. Lorek, Bruce C. Branson, and Rhoda C. Icerman	1992	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV	ratio analysis			time series regression, ARIMA, Box Jenkins, Random Walk, Random Walk Drift	
87	Default on debt obligations and the issuance of going-concern opinions.	Kevin C.W. Chen and Bryan K. Church	1992	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV,R	sampling, ratio analysis			log regression	descriptive statistics
88	Variance Augmentation to Achieve Nominal Coverage Probability in Sampling from Audit Populations	Kermit John Rohrbach	1993	Auditing: A Journal of Practice & Theory	analytical	ST	ratio analysis		Bayesian Theory/Bayesian Belief Networks		
89	An Approach for Calculating Probabilities for Tests of Controls Using an Electronic Spreadsheet	John P. Wendell	1993	Auditing: A Journal of Practice & Theory	discussion	ST	sampling		Bayesian Theory/Bayesian Belief Networks, Naïve Bayes	multivariate distribution, hypergeometric	

										distribution	
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
90	Aggregation of evidence in auditing: A likelihood perspective	Saurav K. Dutta and Rajendra P. Srivastava	1993	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV	ratio analysis		Bayesian Theory/Bayesian Belief Networks, probability theory		
91	The Resolution of Auditor Going Concern Opnions	George E. Nogler	1995	Auditing: A Journal of Practice & Theory	analytical	REV,R	ratio analysis				
92	A Constrained Cost-Minimization Model for Audit Sample Planning	Jack C. Robertson	1995	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	sampling, ratio analysis, firm developed propriet				multicriteria decision aid

							ary software				
93	On the use of Bayesian statistics in the audit process	James C. Loebbecke	1995	Auditing: A Journal of Practice & Theory	conceptual	P,ST,REV	sampling		Bayesian Theory/Bayesian Belief Networks		
94	An Empirical Study of Belief-Based and Probablity-Based Specifications of Audit Risk	Richard Dusenbury, jane L. Reimers and Stephen Wheeler	1996	Auditing: A Journal of Practice & Theory	experimental	P,ST	ratio analysis, firm developed proprietary software		Bayesian Theory/Bayesian Belief Networks		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

95	"REAL-IZING" TIIE BENEFITS OF NEW TECHNOLOGIES AS A SOURCE OF AUDIT EVIDENCE: AN INTERPRETIVE FIELD STUDY	MICHAEL J. FISCHER	1996	Accounting, Organizations and Society	Behavioral	P,ST	CAATS, ratio analysis, firm developed proprietary software				
96	Assessing the risk of management fraud through neural network technology	Bian Patrick Green and Jae Hwa Choi	1997	Auditing: A Journal of Practice & Theory	analytical	P,REV	ratio analysis		ANN		
97	A computational model of loan loss judgments.	William F. Wright and John J. Willingham	1997	Auditing: A Journal of Practice & Theory	behavioral	ST	ratio analysis, firm developed proprietary software		expert systems/decision aids		
98	Sample size determination using the augmented variance estimator	Kermit John Rohrbach	1997	Auditing: A Journal of Practice & Theory	theoretical	ST	sampling				Monte Carlo Simulation

99	Reliance on decision aids: An examination of auditors' assessment of management fraud.	Martha M. Eining, Donald R. Jones and James K. Loebbecke	1997	Auditing: A Journal of Practice & Theory	experimental	P	ratio analysis, firm developed proprietary software		expert systems/decision aids	log regression, step-wise logistic	
100	Stabilising the sieve sample size using PPS	James M. Horgan	1997	Auditing: A Journal of Practice & Theory	theoretical	ST	sampling				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
101	Early Efforts of The US Public Accounting Profession to Investigate The Use Of Statistical Sampling	James j. Tucker, III and Frank C. Lordi	1997	The Accounting Historians Journal	archival		sampling, ratio analysis				

102	An empirical investigation of the relationship between the computerization of accounting systems and the incidence and size of audit differences	Timothy B. Bell, W. Robert Knechel, Jeff L. Payne and John J. Willingham	1998	Auditing: A Journal of Practice & Theory	behavioral	P,ST	sampling, ratio analysis				
103	Detecting and estimating misstatement in two-step sequential sampling with probability proportional to size	Neal B. Hitzig	1998	Auditing: A Journal of Practice & Theory	theoretical	ST	sampling				Monte Carlo Simulation
104	The error detection of structural analytical procedures: A simulation study	Yining Chen and Robert A. Leitch	1998	Auditing: A Journal of Practice & Theory	Theoretical	P,ST,REV	firm developed proprietary software		Simulation	ARIMA, X-11, Martingale	Structural Model
105	Neural network detection of management fraud using published financial data	Kurt Fanning and Kenneth O. Cogger	1998	International Journal of Intelligent Systems in Accounting, Finance & Management	theoretical	P,ST,REV	ratio analysis, firm developed proprietary software		ANN		

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
106	An analysis of the relative power characteristics of analytical procedures	Yining Chen and Robert A. Leitch	1999	Auditing: A Journal of Practice & Theory	analytical	ST	firm developed proprietary software			log regression, step-wise logistic, time series regression, ARIMA, X-11, Martingale	
107	A comparative evaluation of belief revision models in auditing	Ganesh Krishnmoorthy, Theodore J. Mock, and Mary T. Washington	1999	Auditing: A Journal of Practice & Theory	experimental	P,ST,REV,R			Bayesian Theory/Bayesian Belief Networks		

108	The role of audit technology and extension of audit procedures in strategic auditing	Ashutosh Deshmukh	1999	International Journal of Applied Quality Management	theoretical	ST,REV,R	sampling, ratio analysis		probability theory		
109	Error projection and uncertainty in the evaluation of aggregate error	David Burgstahler, Steven M. Glover, and James Jiambalvo	2000	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	sampling, ratio analysis				
110	Decision processes in audit evidential planning: A multistage investigation	Arnold M. Wright and Jean C. Bedard	2000	Auditing: A Journal of Practice & Theory	behavioral	P	sampling, ratio analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

111	Attribute sampling: A belief-function approach to statistical audit evidence	Pter R. Gillett and Rajandra Srivastava	2000	Auditing: A Journal of Practice & Theory	theoretical	ST	sampling		Bayesian Theory/Bayesian Belief Networks		
112	A decision aid for assessing the likelihood of fraudulent financial reporting	Timothy B. Bell and Joseph Carcello	2000	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV	firm developed proprietary software		expert systems/decision aids	log regression	multicriteria decision aid
113	Analytical procedures and audit-planning decisions	Steven M. Glover, James Jiambalvo, and Jane Kennedy	2000	Auditing: A Journal of Practice & Theory	experimental	P,ST,REV	sampling, ratio analysis			log regression	
114	The audit risk model: An empirical test for conditional dependencies among assessed component risks	Richard B. Dusenbury, Jane L. Reimers and Stephen W. Wheeler	2000	Auditing: A Journal of Practice & Theory	experimental	P,ST,REV	sampling, ratio analysis			log regression	
115	An experimental assessment of recent professional developments in nonstatistical audit sampling guidance	William F. Messier, Steven J. Kachelmeier, and Kevan L. Jensen	2001	Auditing: A Journal of Practice & Theory	behavioral	ST	sampling				multicriteria decision aid

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
116	Continuous auditing: Building automated auditing capability.	Zabihollah Rezaee, Ahmad Sharbatoghlie, Rick Elam, and Peter L. McMickle	2002	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV	sampling, firm developed proprietary software, , transaction tests, data modeling, data analytics	clustering, text mining, Visualization		log regression, step-wise logistic, linear regression, time series regression, discriminant analysis,, multivariate regression analysis	Benford's Law, descriptive statistics, multicriteria decision aid

117	The effects of decision aid orientation on risk factor identification and audit test planning	Jean C. Bedard and Lynford E. Graham	2002	Auditing: A Journal of Practice & Theory	behavioral	P,ST	sampling, ratio analysis, firm developed proprietary software				multicriteria decision aid
118	The effectiveness of expectation models in recognizing error patterns and generating and eliminating hypotheses while conducting analytical procedures	Robert A. leitch and Yining Chen	2003	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV				step-wise logistic, ARIMA, Martingale,	Structural Model
119	Applying digital analysis using Benford's law to detect fraud: the dangers of type I errors	Richard Cleary and Jay C. Thibodeau	2005	Auditing: A Journal of Practice & Theory	theoretical	ST	data analytics				Benford's Law
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning,**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

						ST = substantive testing, REV = review, R = report					
120	Peer-based approach for analytical procedures	Rani Hoitash, Alexander Kogan, and Miklos A. Vasarhelyi	2006	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV				time series regression	
121	Auditor specialization, perceived audit quality, and audit fees in the local government audit market	Lowensohn, S., Johnson, L. E., Elder, R. J., & Davies, S. P	2007	Journal of Accounting and Public Policy,	behavioral	E, P, ST, REV, R				Ordinal regression model, linear regression	descriptive statistics
122	Risk monitoring and control in audit firms: A research synthesis	Jean C. Bedard, Donald D. Deis, Mary B. Curtis, and J. Gregory Jenkins	2008	Auditing: A Journal of Practice & Theory	conceptual	E	firm developed proprietary software				multicriteria decision aid

123	Auditors' Cross-Sectional and Temporal Analysis of Account Relations in Identifying Financial Statement Misstatements.	Scott D. Vandervelde, Yining Chen, and Robert A. Leitch	2008	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV				step-wise logistic, Martingale	
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
124	Financial statement fraud: Insights from the academic literature	Chris E. Hogan, Zabihollah Razaee, Richard A. Riley, Jr., and Uma K. Velury	2008	Auditing: A Journal of Practice & Theory	literature review	P,ST,REV	ratio analysis, firm developed proprietary software		ANN, expert systems/decision aids	log regression, linear regression, time series regression, discriminant analysis, multivariate	multicriteria decision aid, Benford's Law, descriptive statistics

Paper Number	Title	Authors	Year	Journal	Paper Category	Stage(s) of Audit Cycle: E = engageme	Techniques: Audit Examinations	Techniques: Unsupervised	Techniques: Supervised	Techniques Regression	Techniques: Other Statistics
										distribution, multivariate regression analysis statistics	
125	The impact of positive and negative mood on the hypothesis generation and ethical judgments of auditors	Anna M. Cianci	2009	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	sampling, ratio analysis				
126	Data diagnostics using second-order tests of Benford's law	Mak J. Nigrini and Steven J. Miller	2009	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV	sampling,				
127	Modified sieve sampling: A method for single-and multi-stage probability-proportional-to-size sampling	Lucas A. Hoogduin, Thomas W. Hall, and Jeffrey J. Tsay	2010	Auditing: A Journal of Practice & Theory	theoretical	ST	sampling				

						nt, P = planning, ST = substantive testing, REV = review, R = report					
128	Audit risk assessments using belief versus probability	Hionori Fukukawa and Theodore J. Mock	2011	Auditing: A Journal of Practice & Theory	experimental	P	ratio analysis		Dempster Shafer Theory, probability theory		
129	Fraudulent management explanations and the impact of alternative presentations of client business evidence	William F. Wright and Leslie Berger	2011	Auditing: A Journal of Practice & Theory	experimental	P,ST,REV					
130	How do audit seniors respond to heightened fraud risk?	Jacqueline S. Hammersly, Karla M. Johnstone, and Kathyrn Kadous	2011	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	sampling, ratio analysis				

131	A review and model of auditor judgments in fraud-related planning tasks	Jackeline S. Hammersly	2011	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV	sampling, ratio analysis, firm developed proprietary software				multicriteria decision aid
132	Strategic analysis and auditor risk judgments	Natalia Kochetova-Kozloski and William Messier, Jr.	2011	Auditing: A Journal of Practice & Theory	experimental	P,ST	sampling, ratio analysis, firm developed proprietary software				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

133	Assessing risk with analytical procedures: Do systems-thinking tools help auditors focus on diagnostic patterns?.	Ed O'Donnell and Jon D. Perkins	2011	Auditing: A Journal of Practice & Theory	experimental	P,ST,REV	sampling, ratio analysis	Visualization			
134	Extreme estimation uncertainty in fair value estimates: Implications for audit assurance.	Brant E. Christiansen, Steven M. Glover, and David A. Wood	2012	Auditing: A Journal of Practice & Theory	conceptual	P,ST,REV	ratio analysis				
135	The impact of initial information ambiguity on the accuracy of analytical review judgments.	Benjamin L. Luipold and Thomas E. Kida	2012	Auditing: A Journal of Practice & Theory	experimental	P,REV	ratio analysis				
136	Two decades of behavioral research on analytical procedures: What have we learned?	Milliam F. Messier, Jr., Chad A. Simon, and Jason L. Smith	2013	Auditing: A Journal of Practice & Theory	literature review	P,ST,REV	sampling, ratio analysis, firm developed proprietary software				Benfords Law

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**
137	Selecting Audit Samples Using Benford's Law	Carlos Gomes da Sliva and Pedro M.R. Carreira	2013	Auditing: A Journal of Practice & Theory	theoretical	P	ratio analysis, firm developed proprietary software, data analytics				Benfords Law
138	Field data on accounting error rates and audit sampling	Michael Durney, Randal J. Elder, and Steven M. Glover	2014	Auditing: A Journal of Practice & Theory	field study	ST	sampling				

139	The Use of Business Risk Audit Perspectives by Non-Big 4 Audit Firms	Joost van Buuren, Christopher Koch, Niels van Nieuw Amerongen, and Arnold M. Wright	2014	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	ratio analysis, data analytics				
140	Design and evaluation of a continuous data level auditing system	Alexander Kogan, Michael G. Alles, Miklos A. Vasarhelyi, and Jia Wu	2014	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV				linear regression, time series regression	Structural Model
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
141	Materiality guidance of the major public accounting firms.	Aasund Eilifsen and William F. Messier, Jr.	2015	Auditing: A Journal of Practice & Theory	theoretical	P,ST,REV	ratio analysis				

142	Further evidence on the auditor's going-concern report: The influence of management plans	Bruce K. Behn, Stephen E. Kaplan, and Kip Krumwiede	2001	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV,R	sampling, ratio analysis			log regression, multivariate regression analysis	
143	The Effect of Domain-Specific Experience on Evaluation of Management Representations in Analytical Procedures	Jean C. Bedard and Stanley F. Biggs	1991	Auditing: A Journal of Practice & Theory	behavioral	ST,REV,R	sampling, ratio analysis				
144	Audit committee composition and auditor reporting	Joseph V. Carcello and Terry L. Neal	2000	Accounting Review	analytical	P,ST,REV,R	sampling, ratio analysis,			log regression, multivariate regression analysis	descriptive statistics
145	Going-concern opinions: The effects of partner compensation plans and client size	Joseph V. Carcello, Dana Hermanson, and H. Fenwick Huss	2000	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV,R	sampling, ratio analysis			log regression	descriptive statistics
146	Audit partner tenure and audit quality	Peter Carey and Roger Simnett	2006	Accounting Review	analytical	P,ST,REV,R	sampling, ratio analysis			log regression	

Paper Number	Title	Authors	Year	Journal	Paper Category	Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report	Techniques: Audit Examinations	Techniques: Unsupervised	Techniques: Supervised	Techniques Regression	Techniques: Other Statistics
147	The effects of output interference on analytical procedures judgments.	John C. Anderson, Steven E. Kaplan, and Philip M. J. Reckers	1992	Auditing: A Journal of Practice & Theory	experimental	P	ratio analysis				
148	Explanation and Counterexplanation during Audit Analytical Review	Koonce, L.	1992	The Accounting Review	behavioral	P	ratio analysis				
149	The endogenous relationship between audit-report type and business termination: Evidence on private firms in	Ann Gaeremynck	2003	Accounting and Business Research	analytical	P,ST,REV,R	sampling, ratio analysis			log regression	

	a non-litigious environment										
150	Bankruptcies, audit reports, and the Reform Act	Marshall A. Geiger and K. Raghunanda n	200 1	Auditing: A Journal of Practice & Theory	analytica l	P,ST,REV ,R	samplin g, ratio analysis			log regression	
151	Statistically incoherent hypothesis tests in auditing.	D.J. Johnstone	199 4	working paper	theoretic al		samplin g		Bayesian Theory/Bayesi an Belief Networks, probability theory		
152	Indicators of exceptions to the going concern assumption	Alan S. Levitan and James A. Knoblett	198 5	Auditing: A Journal of Practice & Theory	analytica l	P,REV	samplin g, ratio analysis			multivaria te distributio n	
Pape r Num ber	**Title**	**Authors**	**Yea r**	**Journal**	**Paper Categor y**	**Stage(s) of Audit Cycle: E = engageme nt, P = planning, ST = substanti ve testing, REV =**	**Techniq ues: Audit Examin ations**	**Techniqu es: Unsuperv ised**	**Techniques: Supervised**	**Techniqu es Regressio n**	**Techni ques: Other Statist ics**

						review, R = report					
153	An analytical investigation of audit qualification decisions in the presence of going concern uncertainties	Krishnagopal Menon and Kenneth B. Schwarz	1987	Contemporary Accounting Research	analytical	P,ST,REV,R	sampling, ratio analysis			log regression	
154	A multivariate analysis of the auditor's going-concern opinion decision.	Jane F. Mutchler	1985	Journal of Accounting Research	analytical	P,ST,REV,R	sampling, ratio analysis			log regression, multivariate regression analysis	
155	The predictive power of limited information in preliminary analytical review: An analytical study	William R. Kinney, Jr.	1979	Journal of Accounting Research	analytical	P,ST,REV,R	ratio analysis, data analytics				descriptive statistics

156	Regression analysis as a means of determining audit sample size	Edward B. Deakin and Michael H. Granof	1974	Accounting Review	theoretical	P,ST,REV,R			Bayesian Theory/Bayesian Belief Networks	log regression, step-wise logistic, univariate regression analysis	
157	Relating statistical sampling to audit objectives	Robert K. Elliott and John R. Rogers	1972	Journal of Accountancy	theoretical	P,ST	sampling				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
158	Evidence on alternative means of assessing prior probability distributions for audit decision making	William L. Felix Jr and Marcia S. Niles	1976	Accounting Review	theoretical	P,ST,REV			Bayesian Theory/Bayesian Belief Networks		

159	A sampling model for audit tests of composite accounts	William L. Felix Jr and Richard A. Grimlund	1977	Journal of Accounting Research	theoretical	ST			Bayesian Theory/Bayesian Belief Networks		Monte Carlo Simulation
160	Research in the auditor's opinion formulation process: State of the art	William L. Felix Jr and William R. Kinney Jr	1982	Accounting Review	literature review	P,ST,REV,R	sampling, ratio analysis				
161	A model for integrating sampling objectives in auditing	Yuji Ijiri and Robert S. Kaplan	1971	Journal of Accounting Research	theoretical	ST	sampling				
162	Assessing prior distributions for applying Bayesian statistics in auditing.	John C Corless	1972	Accounting Review	experimental	ST	sampling		Bayesian Theory/Bayesian Belief Networks		
163	Bayesian statistics in auditing: a comparison of probability elicitation techniques	Michael A. Crosby	1981	Accounting Review	behavioral	ST	sampling		Bayesian Theory/Bayesian Belief Networks		
164	Bayesian statistical methods in auditing	John A. Tracy	1969	Accounting Review	theoretical	ST	sampling		Bayesian Theory/Bayesian Belief Networks		

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**
165	Number magic, auditing acid and materiality- A challenge for auditing research	Ward Edwards	1995	Auditing: A Journal of Practice & Theory	Theoretical		sampling		Bayesian Theory/Bayesian Belief Networks, probability theory		
166	A decision-theory approach to the sampling problem in auditing	William R. Kinney, Jr.	1975	Journal of Accounting Research	theoretical	ST	sampling				
167	The real risks in audit sampling	A.D. Teitlebaum and C.F. Robinson	1975	Journal of Accounting Research	theoretical	P,ST	sampling				
168	A statistical technique for analytical review	Kenneth W. Stringer	1975	Journal of Accounting Research	conceptual	P,REV	firm developed proprietary software			log regression, step-wise logistic, time series regression	

										, multivariate regression analysis	
169	A statistical interpretation of SAP No. 54	Donald M. Roberts	1974	Journal of Accountancy	theoretical	ST	sampling				
170	Auditor's use of analytical review in audit program design	Stanley F. Biggs and Theodore J. Mock	1988	Accounting Review	experimental	P,ST,REV	sampling, ratio analysis				
171	Auditing fair value measurements: A synthesis of relevant research.	Roger D. Martin, Jay S. Rich, and T. Jeffrey Wilks	2006	Accounting Horizons	literature review	P,ST	ratio analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

172	On the use of index models in analytical reviews by auditor	Baruch Lev	1980	Journal of Accounting Research	analytical	ST				linear regression, Sub-Martingale	
173	Human information processing research in accounting: The state of the art in 1982	Robert Libby and Barry Lewis	1982	Accounting, Organizations and Society	literature review	P,ST,REV,R	sampling, ratio analysis		Boosting, Bayesian Theory/Bayesian Belief Networks, probability theory	log regression, linear regression, time series regression, ARIMA, discriminant analysis, multivariate regression analysis	multicriteria decision aid, descriptive statistics
174	Dollar unit sampling: Multinomial bounds for total overstatement and understatement errors.	John Neter, Robert A Leitch and Stephen E Fienberg	1978	Accounting Review	analytical	ST	sampling, ratio analysis				
175	An analytical study of error characteristics in audit populations	John G. Ramage, Abba M. Krieger and Leslie L. Spero	1979	Journal of Accounting Research	analytical	ST	sampling, ratio analysis				

176	A Bayesian approach to asset valuation and audit size	William R. Scott	1973	Journal of Accounting Research	analytical	P,ST,REV,R	sampling, ratio analysis		Bayesian Theory/Bayesian Belief Networks, probability theory		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
177	Employing generalized audit software in the financial services sector: Challenges and opportunities.	Roger Debreceny, Sook-Leng LEE, Willy NEO, and Jocelyn TOH Shuling	2005	Managerial Auditing Journal	behavioral	P,ST,REV	sampling, ratio analysis, firm developed proprietary software, transaction tests, data modeling, data analytics	Visualization	expert systems/decision aids		

178	A stochastic model of the internal control system.	Seongjae Yu and John Neter	1973	Journal of Accounting Research	theoretical	P,ST	sampling, ratio analysis		probability theory		
179	The Four Roles of Sampling in Auditing: Representative, Corrective, Protective, and Preventive (No. RR-165).	Yuji Ijiri and Robert S. Kaplan	1969	CARNEGIE-MELLON UNIV PITTSBURGH PA MANAGEMENT SCIENCES RESEARCH GROUP	theoretical		sampling, ratio analysis				
180	A Review of Developments in Statistical Sampling for Accountants	Lawrence L. Vance	1960	Accounting Review	analytical		sampling, ratio analysis				
181	Bayesian statistics: A review	J. G. Birnberg	1964	Journal of Accounting Research	analytical		sampling, ratio analysis		probability theory		
182	Bayesian analysis in auditing	James E. Sorenson	1969	Accounting Review	theoretical		sampling, ratio analysis		Bayesian Theory/Bayesian Belief Networks, probability theory		

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
183	The effectiveness of statistical analytical review on overall audit effectiveness: A simulation analysis	W. Robert Knechel	1988	Accounting Review	analytical	P,ST,REV	sampling, ratio analysis			time series regression, univariate regression analysis, multivariate regression analysis	
184	Implications of prior probability elicitation on auditor sample size decisions	Michael A. Crosby	1980	Journal of Accounting Research	behavioral	P,ST	sampling, ratio analysis		Bayesian Theory/Bayesian Belief Networks, probability theory		
185	On Combining Evidence from Subpopulations into a Composite Conclusion	J. EFRIM BORI1Z PING ZHANG STEVE	1993	Contemporary Accounting Research	empirical	E,P,REV	sampling		Expert Systems/Decision Aids, Dempster Shafer Theory,		

		ALDERSL EY							Probability Theory		
186	The impact of the review process in hypothesis generation tasks	Zubaidah Ismail and Ken T. Trotman	199 5	Accounting , Organizatio ns and Society	behavior al	REV	ratio analysis				
187	A statistical paradox in auditing.	D.J. Johnstone	199 4	Abacus	conceptu al	ST	samplin g		Bayesian Theory/Bayesi an Belief Networks, probability theory		
Pape r Num ber	**Title**	**Authors**	**Yea r**	**Journal**	**Paper Categor y**	**Stage(s) of Audit Cycle: E = engageme nt, P = planning, ST = substanti ve testing, REV = review, R = report**	**Techniq ues: Audit Examin ations**	**Techniqu es: Unsuperv ised**	**Techniques: Supervised**	**Techniqu es Regressio n**	**Techni ques: Other Statist ics**

188	A descriptive analysis of computer audit specialists decision-making behavior in advanced computer environments.	Stanley F. Biggs, William F. Messier, Jr., and James V. Hansen	1987	Auditing: A Journal of Practice & Theory	behavioral	ST	ratio analysis, firm developed proprietary software, transaction tests, data modeling, data analytics		expert systems/decision aids, probability theory		
189	Statistics Offers a Solution to Tomorrow's Auditing Complexities	Robert C. Mogis and Donald Rogoff	1962	Accounting Review	normative	ST	sampling				
190	Are Industry Specialists More Efficient and Effective in Performing Analytical Procedures? A Multi-stage Analysis	Wendy Green	2008	International Journal of Auditing	behavioral	P,ST,REV	ratio analysis				

191	Bayesian sampling procedures for auditors: computer-assisted instruction	Edward Blocher and Jack C. Robertson	1976	Accounting Review	education case study		sampling		Bayesian Theory/Bayesian Belief Networks, probability theory		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
192	Subjective Probability Elicitation: The Effect of Congruity of Datum and Response Mode on Performance	G.R. Chesley	1977	Journal of Accounting Research	theoretical	P,REV	ratio analysis		probability theory		
193	Testing the Consistency of Auditors' Prior Distributions and Sampling Results	Donald R. Nichols and R.C. Baker	1977	Abacus	theoretical	P,ST,REV,R	sampling, ratio analysis		Bayesian Theory/Bayesian Belief Networks, probability theory		

194	Statistical sampling in auditing	J.M. Sully	1974	Journal of the Royal Statistical Society	theoretical	P,ST,REV	sampling, ratio analysis				
195	The relationship of internal control evaluation and audit sample size	Kenneth A. Smith	1972	Accounting Review	theoretical	ST	sampling		Bayesian Theory/Bayesian Belief Networks, probability theory		
196	Auditor decision making on overall system reliability: accuracy, consensus, and the usefulness of a simulation decision aid	Ron Weber	1978	Journal of Accounting Research	behavioral	P,ST	ratio analysis		expert systems/decision aids		
197	Subjective prior probability distributions and audit risk	Paul J. Beck, Ira Solomon, and Lawrence A. Tomassini	1985	Journal of Accounting research	theoretical	P,ST	sampling, ratio analysis		Bayesian Theory/Bayesian Belief Networks		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substanti**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

						ve testing, REV = review, R = report					
198	Initial Efforts of Kenneth W. Stringer to Develop a Statistical Sampling Plan	James J. Tucker, III	199 4	The Accounting History Journal	archival	P,ST	samplin g				
199	A note on a comparison of Bayesian with non-Bayesian dollar-unit sampling bounds for overstatement errors of accounting populations	Wally Smieliauska s	198 6	Accounting Review	theoretic al	ST	samplin g		Bayesian Theory/Bayesi an Belief Networks, probability theory		
200	A mathematical approach to the analysis and design of internal control systems	Barry E. Cushing	197 4	Accounting Review	theoretic al	P,ST	samplin g, ratio analysis		probability theory		
201	A Note on the Interface Between Compliance and Substantive Tests	Andrew D. Bailey, Jr., and Daniel L. Jensen	197 7	Journal of Accounting Research	theoretic al	P,ST			Bayesian Theory/Bayesi an Belief Networks, probability theory		

202	A finite population bayesian model for compliance testing.	James T. Godfrey and Richard W. Andrews	1982	Journal of Accounting Research	theoretical	P,ST	sampling		Bayesian Theory/Bayesian Belief Networks, probability theory		
203	Decision theory aspects of internal control system design/compliance and substantive tests	William R. Kinney, Jr.	1975	Journal of Accounting Research	theoretical	P,ST,REV	sampling, ratio analysis		expert systems/decision aids		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
204	A synthesis of fraud-related research	Gregory M. Trompeter, Tina D. Carpenter, Naman Desai, Keith L. Jones, and Richard A. Riley, Jr.	2013	Auditing: A Journal of Practice & Theory	literature review	P,ST,REV	traditional analytcical procedures				

205	The audit of fair values and other estimates: The effects of underlying environmental, task, and auditor-specific factors.	Brian Batten, Lisa Milici Gaynor, Linda McDaniel, Norma R. Mantague, and Gregory E. Sierra	2013	Auditing: A Journal of Practice & Theory	literature review	P,ST	ratio analysis		expert systems/decision aids		
206	Audit reporting for going-concern uncertainty: A research synthesis	Elizabeth Carson, Neil L. Fargher, Marshall A. Geiger, Clive S. Lennox, K. Raghunandan, and Marleen Willekens	2013	Auditing: A Journal of Practice & Theory	literature review	REV,R	ratio analysis				
207	FSA: Applying AI techniques to the familiarization phase of financial decision making.	Chunka Mui abd William E. McCarthy	1987	IEEE Expert	theoretical	P,ST	ratio analysis, firm developed proprietary software		expert systems/decision aids		

Paper Number	Title	Authors	Year	Journal	Paper Category	Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report	Techniques: Audit Examinations	Techniques: Unsupervised	Techniques: Supervised	TechniquesRegression	Techniques: Other Statistics
208	The Expert Systems and Their Role in Developing External Auditor's Performance and Improving Audit Service's Quality in Information Technology Environment in Audit's Offices Located in the Hashemite Kingdom of Jordan	Reem Okab	2013	International Journal of Business and Management	behavioral	P,ST,REV	firm developed proprietary software		expert systems/decision aids		

209	Using electronic audit workpaper systems in audit practice: Task analysis, learning, and resistance	Jean C. Bedard, Michael L. Ettridge, and Karla M. Johnstone	2006	Learning and Resistance	behavioral	P,ST,REV	ratio analysis, firm developed proprietary software		expert systems/decision aids		
210	The effect of electronic audit environments on performance	Lynn Bible, Lynford Graham, and Andrew Rosman	2005	Journal of Accounting, Auditing, and Finance	behavioral	E,P,ST,REV	ratio analysis, firm developed proprietary software		expert systems/decision aids		
211	An Empirical Investigation of the Auditor's Decision to Project Errors	Randal J. Elder and Robert D. Allen	1998	Auditing: A Journal of Practice & Theory	empirical	ST	Sampling				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

212	Auditors' assessment and incorporation of expectation precision in evidential analytical procedures	Linda S. McDaniel and Laura E. Simmons	2007	Auditing: A Journal of Practice & Theory	experimental	P,ST,REV	sampling, ratio analysis				
213	Independent Auditors' Responsibilties for Violations of Anti-bribery Provisions Under the U.S. Foreign Corrupt Practices Act: Auditing for Bribes	Ricardo Colono	2015	Journal of Forensic and Investigative Accounting	conceptual	E,P,ST,REV,R	sampling, ratio analysis, transaction tests, data modeling, data analytics	text mining, Visualization			
214	Using Pattern Analysis Methods to Supplement Attention Directing Analytical Procedures	James R. Coakley	1995	Expert Systems with Applications	empirical	P	ratio analysis		ANN		
215	Using CAATTs in preliminary analytical review to enhance the auditor's risk assessment	Alex Vuchnich	2008	CPA Journal	conceptual	P	ratio analysis, transaction tests, data modeling, data analytics	Visualization		log regression, linear regression, time series regression, multivaria	

										te distribution, multivariate regression analysis	
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
216	Continuous Online Auditing: A program of research	Alexander Kogan Ephraim F. Sudit Miklos A. Vasarhelyi	1999	Journal of Information Systems	conceptual	P,ST,REV	CAATS, sampling, ratio analysis, firm developed proprietary software	Visualization	expert systems/decision aids	time series regression	Benford's Law

217	Identifying audit adjustments with attention-directing procedures	Arnold Wright and Robert H. Ashton	1989	Accounting Review	survey	P, ST, R	sampling, ratio anlysis			log regression, linear rehression, time series regression	
218	Glossary of Statistical Terms for Accountants and Bibliography on the Application of Statistical Methods to Accounting, Auditing, and Management Control	American Institute of Certified Public Accountants	1958	white paper	literature review		sampling, ratio analysis				
219	AN EMPIRICAL INVESTIGATION OF THE USE OF ANALYTICAL REVIEW BY EXTERNAL AUDITORS	IAN A. M. FRASER DAVID J. HATHERLY KENNY Z. LIN	1997	The British Accounting review	behavioral survey	P,ST,REV	ratio analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P =**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**

						planning, ST = substantive testing, REV = review, R = report					
220	Empirical Analyses of Errors Discovered in Audits of Property and Casualty Insurers	Timothy Bell and W. Robert Knechel	1994	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	sampling, ratio analysis				
221	Between a Rock and a Hard Place: A Path Forward for Using Substantive Analytical Procedures in Auditing Large P&L Accounts Commentary and Analysis	Steven M. Glover, Douglas F. Prawitt, and Michael S. Drake	2015	Auditing: A Journal of Practice & Theory	analytical	P, ST, REV, R	sampling, ratio analysis			log regression, linear regression	
222	An analytical study of regression analysis as an analytical procedure	WILSON, A. C., & HUDSON, D	1989	Contemporary Accounting Research	analytical	E,ST,REV				log regression, single exponential smoothing model, double exponenti	multicriteria decision aid, descriptive statistics

										al smoothing model, univariate regression analysis, multivariate regression analysis	
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
223	Auditors' Experience with Material Irregularities: Frequency , Nature , and Detectability	James K. Loebbecke , Martha M. Eining, and John J. Willingham	1989	Auditing: A Journal of Practice & Theory	behavioral	P,ST,REV	Sampling, Ratio Analysis , Firm developed proprietary software				

224	ATTENTION-DIRECTING ANALYTICAL REVIEW USING ACCOUNTING RATIOS-A CASE-STUDY	William R. Kinney	1987	Auditing: A Journal of Practice & Theory	case study	P, R	ratio analysis				
225	Behind the Numbers: Insights into Large Audit Firm Sampling Policies	Brant E. Christiansen, Randal J. Elder, and David A. Wood	2014	Accounting Horizons	survey	ST	transaction tests, data analytics, sampling, firm developed proprietary software				
226	Audit sampling research: A synthesis and implications for future research	Randal J. Elder, Abraham D. Akresh, Steven M. Glover, Julia L. Higgs, and Jonathan Liljegren	2013	Auditing: A Journal of Practice & Theory	literature review	P,ST,R	sampling				

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
227	Consequences of Big Data and Formalization on Accounting and Auditing Standards	John Peter Krahel and William R. Titera	2015	Accounting Horizons	conceptual	P,ST	sampling, ratio analysis	process mining, clustering, fuzzy logic, text mining	random forest, BBN,	log regression, linear regression, time series regression	descriptive statistics
228	Big Data Analytics in Financial Statement Audits	Min Cao, Roman Chychyla, and Trevor Stewart	2015	Accounting Horizons	conceptual	E,P,ST,REV,R	ssampling, ratio analysis	clustering, text mining			
229	How Big Data Will Change Accounting	J. Donald Warren Jr., Kevin C. Moffitt, and Paul Byrnes	2015	Accounting Horizons	conceptual	ST	ratio analysis	clustering, text mining, visualization			descriptive statistics

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**
230	Technology in audit engagements: a case study	Miklos A. Vasarhelyi and Silvia Romero	2014	Managerial Auditing Journal	behavioral field study	E,P,ST,REV,R	transaction tests, data modeling, data analytics, sampling, ratio analysis, firm developed proprietary software		expert systems/decision aids,		
231	A fuzzy neural network for assessing the risk of fraudulent financial reporting	Jerry W. Lin, Mark I. Hwang, Jack D. Becker	2003	Managerial Auditing Journal	empirical	ST	ratio analysis	clustering, fuzzy logic	ANN, Fuzzy ANN, expert systems/decision support aids	log regression	

232	A Comparison of Selected Artificial Neural Networks that Help Auditors Evaluate Client Financial Viability	Harlan L. Etheridge, Ram S. Sriram, H. Y. Kathy Hsu	2000	Decision Sciences	empirical	REV,R	ratio analysis		ANN, Fuzzy ANN, Expert systems/Decision aids, BBN	log regression	multivariate distribution, multivariate regression analysis
233	Artificial Neural Network Models for Predicting Patterns in Auditing Monthly Balances	E. Koskivaara	2000	The Journal of Operational Research Society	empirical	P,ST	ratio analysis		ANN, Fuzzy ANN	log regression, linear regression	
234	Comparison of Analysis Performed By Classical Approach and Bayesian Approach in Auditors' Decision Making Process	Nurten Erdogana, Sezen Uludagb	2014	Procedia - Social and Behavioral Sciences	behavioral	E,P,ST,REV,R			BBN, Bayesian Structural Time Series	log regression	

235	Financial statement error: client's business risk assessment and auditor's substantive test	Halil Paino a*, Khairul Anuar Abdul Hadi a and Wan Mardyatul Miza Wan Tahir	2014	Procedia - Social and Behavioral Sciences	behavioral	P,ST	transaction tests, atio analysis				
236	The case for process mining in auditing: Sources of value added and areas of application	Mieke Jans a, Michael Alles b,∗, Miklos Vasarhelyi	2013	International Journal of Accounting Information Systems	conceptual	P.ST.REV	transaction tests, ratio analysis	process mining, clustering			
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**TechniquesRegression**	**Techniques: Other Statistics**
237	Data Analytics for Financial Statement Audits	Trevor R. Stewart	2015	AICPA publication, "Audit Analytics"	conceptual	E,P,ST,REV,R	CAATS, sampling, ratio analysis, firm developed	process mining, clustering, text mining, visualization	ANN, expert systems/decision aids	log regression, linear regression, time series regression	multivariate regression analysis

							proprietary software				
238	The Impact of Hypothesis Set Size on the Time Efficiency and Accuracy of Analytical Review Judgments	Sudip Bhattacharjee, Thomas Kida and Dennis M. Hanno	1999	Journal of Accounting Research	behavioral	P,ST	Ratio Analysis				
239	An Exploratory Study of Auditors' Evidential Planning Judgments	Theodore J. Mock and Arnold Wright	1993	Auditing: A Journal of Practice & Theory	empirical	P,ST	Sampling				
240	An Archival Study of the Existence, Cause, and Discovery of lncome-Affecting Financial Statement Misstatements	GARY EN1WISTLE DARYL LINDSAY	1994	Contemporary Accounting Research	behavioral	P, ST	ratio analysis				

241	Financial Statement Errors and Internal Control Judgments	John J. Willingham and William F. Wright	1985	Auditing: A Journal of Practice & Theory	empirical	ST,REV	sampling, ratio analysis				descriptive statistics
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
242	Error Characteristics in Audit Populations: Their Profile and Relationship to Environmental Factors	Richard W. Kreutzf eldt and Wanda A. Wallace	1986	Auditing: A Journal of Practice & Theory	empirical	ST	sampling, ratio analysis				
243	Statistical or Non-statistical Sampling: Which Approach Is Best?	Dr. Janet L. Colbert	1991	Journal of Applied Business Research	conceptual	P,ST	Sampling				

244	An Empirical Study of Error Characteristics in Accounting Populations	Jane Ham, Donna Losell and Wally Smieliauskas	1985	The Accounting Review	empirical	P.ST	Sampling, Ratio Analysis				Descriptive Statistics
245	A Stochastic Model of Error Generation in Accounting Systems	W. Robert Knechel	1985	Accounting and Business Research	empirical	P,ST,REV	Sampling, Ratio Analysis			Linear Regression	
246	Characteristics of Errors in Accounts Receivable and Inventory Audits	Johnny R. Johnson, Robert A. Leitch, and John Neter	1981	The Accounting Review	empirical	P,ST,REV	Sampling, Ratio Analysis			Linear Regression	Descriptive Statistics
247	Towards a Contingency View of Audit Evidence	Arnold Wright and Theodore J. Mock	1985	Auditing: A Journal of Practice & Theory	conceptual	P.ST.REV	Sampling, Ratio Analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

248	Sample Error Characteristics and Projection of Error to Audit Populations	David Burgstahler and James Jiambalvo	198 6	The Accounting Review	behavior al	ST	Samplin g, Ratio Analysis				
249	Error Rates, Error Projection, and Consideration of Sampling Risk: Audit Sampling Data from the Field	Michael Durney Randal Elder Steven Glover	201 3		empirica l	ST	Samplin g				
250	Audit Program Planning Using A Belief Function Framework	Theodore J. Mock, Arnold Wright, and Rajendra P. Srivastava	199 8	Proceedings of the 1998 Deloitte & Touche University of Kansas Symposium on Auditing Problems	behavior al	P			DS Theory, probability theory		
251	A Longitudinal Field Investigation of Auditor Risk Assessments and Sample Size Decisions	Randal J. Elder Robert D. Allen	200 3	The Accounting Review	empirica l	P,ST	Samplin g, Ratio Analysis				
252	Toward a More Consistent Model for Audit Risk	John T. Sennetti	199 0	Auditing: A Journal of Practice & Theory	conceptu al	P,ST	Samplin g, Ratio Analysis		Decision Trees		

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
253	Group Audits, Group-Level Controls, and Component Materiality: How Much Auditing Is Enough?	Trevor R. Stewart William R. Kinney, Jr.	2013	The Accounting Review	empirical	ST	Sampling		BBN, Probability Theory		
254	The Use of Benford's Law as an Aid in Analytical Procedures	Maek J, Nigrini and Linda J. Mittermaier	1997	Auditing: A Journal of Practice & Theory	case study	P,ST,REV	data analytics				Benford's Law
255	The Influence of Sample Characteristics in Sample Evaluation	Edward Blocher and Joseph H. Hylinski	1985	Auditing: A Journal of Practice & Theory	behavioral	ST	Sampling				

256	Application of a Decision Aid in the Judgmental Evaluation of Substantive Test of Details Samples	Stephen A . Butler	1985	Journal of Accounting Research	behavioral	ST	Sampling				
257	Examination of the Effect of Risk Model Components on Perceived Audit Risk	Jerry R. Strawser	1991	Auditing: A Journal of Practice & Theory	behavioral	P	sampling, ratio analysis				
258	Relating Statistical Sampling To Audit Objectives	By Robert K. Ellioit And John R. Rogers	1972	The Journal of Accountancy	conceptual	P,ST	Sampling, Ratio Analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

259	Investigating the Use of Analytical Procedures: An Update and Extension	Elsie C. Ameen and Jerry Strawser	1994	Auditing: A Journal of Practice & Theory	survey	E,P,ST,REV,R	sampling, ratio analysis			log regression, linear regression, time series regression	
260	Strategic Considerations for Unaudited Account Values in Analytical Review	John J. Wild and Stanley F. Biggs	1990	The Accounting Review	empirical	P			Expert Systems/Decision Aids, BBN		
261	The Joint Influence of the Extent and Nature of Audit Evidence, Materiality Thresholds, and Misstatement Type on Achieved Audit Risk	David V. Budescu, Mark E. Peecher, and Ira Solomon	2012	Auditing: A Journal of Practice & Theory	empirical	P	Sampling, Ratio Analysis		Expert System/Decision Aid		
262	Assessing the performance of analytical procedures: a best case scenario	Stephen Wheeler and Kurt Pany	1990	Accounting Review	analytical	P,ST,REV	sampling, ratio analysis			log regression, linear regression, time series regression, Martingale, Sub-Martingale	

263	THE CONTINUOUS AUDIT OF ONLINE SYSTEMS	Miklos A. Vasarhelyi and Fern Halper	1991	Auditing: A Journal of Practice & Theory	conceptual	P,ST,REV	Ratio analysis, Firm developed proprietary software				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
264	Pattern Recognition, Hypotheses Generation, and Auditor Performance in an Analytical Task	Jean C. Bedard and Stanley F. Biggs	1991	The Accounting Review	behavioral	P	ratio analysis				

265	Principles of Analytic Monitoring for Continuous Assurance	Miklos A. Vasarhelyi Michael G Alles Alexander Kogan	2004	Journal of Emerging Technologies in Accounting	conceptual	P,ST	CAATS, sampling, ratio analysis, firm developed proprietary software		Expert Systems/Decision Aids	time series regression	
266	Remote Audit: A Review of Audit-Enhancing Information and Communication Technology Literature	Ryan A. Teeter, Rutgers, Miklos A. Vasarhelyi.	2010	Journal of Emerging Technologies in Accounting	conceptual	ST	CAATS, sampling, ratio analysis, firm developed proprietary software	Process Mining			
267	A SIMULATION STUDY OF THE RELATIVE EFFECTIVENESS OF ALTERNATIVE ANALYTICAL REVIEW PROCEDURES	W. Robert Knechel	1986	Decision Sciences	empirical	ST	ratio analysis			time series, ARIMA, X-11, Martingale	

Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
268	Auditing Firms' Fraud Risk Assessment Practices	Sandra Waller Shelton, O. Ray Whittington, and David Landsittel	2001	Accounting Horizons	behavioral	E, P	ratio analysis		Expert Systems/Decision Aids		
269	Analytical Review Developments in Practice: Misconceptions, Potential Applications, and Field Experience	Wanda A. Wallace	1982		conceptual		sampling, ratio analysis			linear regression, time series regression	structural model
270	A roadmap for future neural networks research in auditing and risk assessment	Thomas G. Calderon, John J. Cheh	2002	Journal of Accounting Information Systems	literature review	P,REV			ANN, General Adaptive ANN, Fuzzy ANN, ID3	log regression	

271	HYPOTIIESIS REVISION STRATEGIES IN CONDUCTING ANALYTICAL PROCEDURES	S. K. ASARE and A. WRIGHT	1997	Accounting, Organizations and Society	behavioral	P,ST,REV	Ratio Analysis				
272	Stein's Paradox and Audit Sampling	Yuji Ijiri and Robert A. Leitch	1980	Journal of Accounting Research	empirical	P,ST			BBN, BSTS		
273	Auditors' Assessments of the Likelihood of Error Explanations in Analytical Review	Vicky B. Heiman	1990	The Accounting Review	behavioral	P, REV	ratio analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

274	Auditing in the e-commerce era	Ning Zhao, David C. Yen and I-Chiu Chang	2004	Information Management & Computer Security	conceptual	P,ST	CAATS, ratio analysis, firm developed proprietary software		Expert Systems/Decision Aids		
275	Can financial ratios detect fraudulent financial reporting?	Kathleen A. Kaminski, T. Sterling Wetzel, and Liming Guan	2004	Managerial Auditing Journal	empirical	P	Ratio Analysis				
276	Why Do Auditors Over-Rely on Weak Analytical Procedures? The Role of Outcome and Precision	Steven M. Glover, Douglas F. Prawitt, and T. Jeffrey Wilks	2005	Auditing: A Journal of Practice & Theory	experimental, behavioral	P,ST,REV	Ratio Analysis				
277	Distributions of Financial Accounting Ratios: Some Empirical Evidence	Edward B. Deakin	1976	The Accounting Review	empirical		Ratio Analysis	Clustering			
278	Updating Audit Standard—Enabling Audit Data Analysis	William R. Titera	2013	Journal of Information Systems	conceptual	E,P,ST,REV,R	Ratio Analysis			Linear Regression	

279	Artificial neural networks in analytical review procedures	Eija Koskivaara	2004	Managerial Auditing Journal	literature review	P,ST,REV	Ratio Analysis		ANN, General Adaptive ANN, Fuzzy ANN	Linear Regression	
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
280	ANALYTICAL PROCEDURES: PROMISE, PROBLEMS AND IMPLICATIONS FOR PRACTICE	STANLEY F. BIGGS, THEODORE J. MOCK and ROGER SIMNETT	1999	Australian Accounting Review	literature review	P,REV	ratio analysis, sampling		decision aids	linear regression, time series regression	
281	Evaluating the sufficiency of causes in audit analytical procedures	Urton Anderson and Lisa Koonce	1998	Auditing: A Journal of Practice & Theory	experimental	P,ST	sampling, ratio analysis				

282	Successful audit workpaper review strategies in electronic environments	Andrew Rosman, Stanley Biggs, Lynford Graham, and Lynn Bible	2007	Journal of Accounting, Finance & Management	behavioral	E,P,ST,REV	ratio analysis, firm developed proprietary software		expert systems/decision aids		
283	Bayesian Fraud Risk Formula for Financial Statement Audits	Rajendra P. Srivastava, Theodore J. Mock, Jerry L. Turner	2008	Abacus	Conceptual	P			Decision Trees, BBN, probability theory		
284	The Dempster-Shafer Theory of Belief Functions for Managing Uncertainties: An Introduction and Fraud Risk Assessment Illustration	Rajendra P. Srivastava, Theodore J. Mock, Lei Gao	2011	Australian Accounting Review	conceptual	P			probability theory, DS Theory		
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV =**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

						review, R = report					
285	A conceptual framework and case studies on audit planning and evaluation given the potential for fraud	Jerry L. Turner Theodore J. Mock Rajendra P. Srivastava	2002	Bond Business School Publications	case study, behavioral	P			BBN, DS Theory		
286	Improving Analytical Procedures: A Case of Using Disaggregate Multilocation Data	Robert D. Allen, Mark S. Beasley, and Bruce Branson	1999	Auditing: A Journal of Practice & Theory	case study	P,REV,R				time series regression, multivariate regression analysis	
287	Analytical Review Procedures	William R. Kinney. J r., and William L. Felix, Jr	1980	Journal of Accountancy	conceptual	P.ST.REV	ratio and trend analysis			time series regression	Structural Model
288	Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the sec*	Dechow, P. M., Sloan, R. G., & Sweeney, A. P.	1996	Contemporary Accounting Research	analytical	E, P				log regression	descriptive statistics

289	Audit Practices of PricewaterhouseCoopers	Barry N. Winograd, James S. Gerson, and Barbara L. Berlin	2000	Auditing: A Journal of Practice & Theory	discussion	E,P,ST,REV,R					
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
290	Multiple hypothesis evaluation in auditing	Rajendra P. Srivastava, Arnold Wright, Theodore J. Mock	2002	Accounting and Finance	conceptual	P, ST			probability theory		

291	The Effect of Autocorrelation on Regression Based Model Efficiency and Effectiveness in Analytical Review	Arlette C. Wilson	1992	Auditing: A Journal of Practice & Theory	analytical	P,ST,REV				log regression, linear regression, time series regression, ARIMA, single exponential smoothing model, double exponential smoothing model	
292	An Investigation of Auditor Judgment in Analytical Review	Stanley F. Biggs and John J. Wild	1985	The Accounting Review	behavioral	P, ST, REV	ratio analysis		decision aids	log regression, linear regression, time series regression	
293	Experience and the Ability to Explain Audit Findings	Robert Libby and David M. Frederick	1990	Journal of Accounting Research	behavioral	P	ratio analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P =**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**

						planning, ST = substantive testing, REV = review, R = report					
294	Judgmental Versus Random Sampling in Auditing: An Experimental Investigation	Lawrence A. Ponemon and John P. Wendell	1995	Auditing: A Journal of Practice & Theory	behavioral	ST	sampling				
295	Multi-Auditor Decision Making on Internal Control System Reliability: A Default Reasoning approach	Ai-Mei Chang, Andrew D. Bailey, Jr., and Andrew B. Whinston	1993	Auditing: A Journal of Practice & Theory	theoretical	ST,REV	data analytics		Bayesian Theory/Bayesian Belief Networks		
296	Inherent Risk	Clarence W. Houghton and John A. Fogerty	1991	Auditing: A Journal of Practice & Theory	survey	P	ratio analysis				
297	Are Auditors' Judgments Sufficiently Regressive?	Edward J. Joyce and Gary C. Biddle	1981	Journal of Accounting Research	behavioral	P, REV			probability theory	linear regression	
298	Anchoring and Adjustment in Probabilistic Inference in Auditing	Edward J. Joyce and Gary C. Biddle	1981	Journal of Accounting Research	behavioral	P, ST			prabability theory, BBN		

299	Mitigating the Consequences of Anchoring in Auditor Judgments	William R. Kinney, Jr. and Wilfred C. Uecker	1982	The Accounting Review	behavioral	P, REV	ratio analysis				
Paper Number	**Title**	**Authors**	**Year**	**Journal**	**Paper Category**	**Stage(s) of Audit Cycle: E = engagement, P = planning, ST = substantive testing, REV = review, R = report**	**Techniques: Audit Examinations**	**Techniques: Unsupervised**	**Techniques: Supervised**	**Techniques Regression**	**Techniques: Other Statistics**
300	Auditing Revenue Account: Can Audit Sampling Replace Substantive Analytical Procedures?	Kyunghee Yoon	2016	dissertation	conceptual	P, ST, REV	sampling			log regression, step wise regression, linear regression, ARIMA, X-11, Sub-Martingale	univariate regression, multivariate regression, strustural model

301	Evidence on the Effect of Financial and Nonfinancial Trends on Analytical Review	Jeffrey R. Cohen, Ganesh Krishnamoorthy, and Arnold M. Wright	2000	Auditing: A Journal of Practice & Theory	experinental	P, ST, REV	sampling, ratio analysis			log regression, linear regression	

Table 23: Some attributes of the 301 papers discussing audit analytics applied in the traditional external audit framework. Phase Six or Continuous Activities does not appear here due to space concerns, given that there were zero papers that discussed analytics in this phase.