Spring 2020

The Effect of Audit Procedure Extent and Precision on Auditors' Skeptical Judgments

Sarah Alexis Judge

THE EFFECT OF AUDIT PROCEDURE EXTENT AND PRECISION ON AUDITORS' SKEPTICAL JUDGMENTS

by

Sarah Alexis Judge

Bachelor of Science
University of South Carolina, 2013

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Business Administration

Darla Moore School of Business

University of South Carolina

2020

Accepted by:

Scott D. Vandervelde, Major Professor

Clark Hampton, Committee Member

David Hitchcock, Committee Member

Chad Stefaniak, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

DEDICATION

To my husband, Billy: Thank you for supporting me, cheering me on, and encouraging me to follow my dream. I could never have done this without you, and I'm grateful for the support that you provided and the sacrifices that you made. We did it!

To my son, William: I hope that when you are older knowing that I followed my dreams will encourage you to follow yours. More importantly, I pray that whatever you do in your life, you use it to honor Jesus and share His love with others.

To my Mom: For as long as I can remember, I wanted to follow in your footsteps. Now that I have, I am even more amazed at everything you have done. Thank you for setting a wonderful example for me.

To my Dad: Thank you for being a constant encouragement in my life. You always reminded me that my worth was not in my education or success, and for that I am grateful.

To my sister, Rachel: Thank you for always being there to talk and listen. I love you behind my back infinity and beyond.

To my brothers, Shawn and Shane: One out of three isn't bad. Thank you for being the best brothers that I could ask for.

To my village in Columbia: Thank you for making Columbia home.

Psalm 62:2 Truly he is my rock and my salvation; he is my fortress; I will never be shaken.

ACKNOWLEDGMENTS

I am grateful to Scott Vandervelde for all of the guidance and support he has provided as my dissertation chair. Thank you Scott for being so generous with your time as you spent countless hours discussing this idea and read countless drafts. I also acknowledge and thank the other members of my committee, Clark Hampton, David Hitchcock, and Chad Stefaniak, for all of their feedback. Having all of you on my committee made this dissertation possible. I am grateful to Amanda Carlson, Lori Shefchik Bhaskar, Paul Black, Scott Jackson, Jared Jones, Ethan LaMothe, Drew Newman, Jason Rasso, Donna Bobek Schmitt, Kristen Steury, Grazia Xiong, and Aaron Zimbelman, along with participants and anonymous reviewers from the 2020 Audit Midyear Conference for their helpful comments and suggestions.

ABSTRACT

I investigate how the extent and precision of an audit procedure influence auditors' responses to errors. The results suggest that while audit procedures with expanded extents are beneficial, this benefit is influenced by the presence of false positives. Specifically, I predict and find that as the extent of an audit procedure expands, auditors will be more likely to recommend audit adjustments. However, this effect is moderated by the presence of false positives. When false positives are absent, auditors are more likely to recommend an adjustment when they use a large extent procedure, however when false positives are present, there is no difference in the likelihood of recommending an adjustment for small and large extent audit procedures. The effect of extent on likelihood judgments is mediated by auditors' reliability judgments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Audit error response is a critical component of the audit process, and auditors are required by standards to revise their risk assessments and planned procedures when a potential misstatement is uncovered (PCAOB 2010a). The current paper aims to investigate how auditors' responses to errors are influenced not only by the numeric value of the discovered error, but also by the attributes of the audit procedure that discovered the error, specifically the extent and precision. The extent of an audit procedure refers to how broadly an account is tested and can range from testing a small sample to testing the entire population, while an audit procedure's precision is defined in this paper as how accurately the procedure can identify true errors.[1] While testing entire populations has not been common practice for an audit due to the high cost (PCAOB 2017), Big Data analytics and the related technology is expected to transform the audit landscape by allowing auditors to test complete populations at a reasonable cost (Cao, Chychyla, and Stewart 2015; EY Reporting 2015; Minniti and Camehl 2018).

Big Data analytics, defined as "the process of inspecting, cleaning, transforming, and modeling Big Data to discover and communicate useful information and patterns,

[1] While not the focus of the current study, precision also diminishes if the audit procedure does not identify errors that are present. This decrease in precision is much more difficult to identify as auditors will likely only learn this after the fact, such as if an error is discovered in the following year which requires a restatement for the previous year. For the current paper, the definition of precision refers to how accurately the procedure discovers true errors.

suggest conclusions, and support decision making," will allow auditors to analyze all of the data available, resulting in complete population testing (Cao, Chychyla, and Stewart 2015, 424). Big data analytics can be used to obtain audit evidence during any stage of the audit, such as identifying anomalies in a substantive test of details (Minnitti and Camehl 2018). It is anticipated that testing the entire population, compared to a sample, will lead auditors to identify a greater number of errors (Krahel and Titera 2015; Vasarhelyi, Kogan, and Tuttle 2015), and will provide greater certainty over the true error amount present in the population. However, expanding the extent to the entire population is expected to diminish the precision of the audit procedure by increasing the identification of false positives (Krahel and Titera, 2015; Vasarhelyi, Kogan, and Tuttle 2015; Yoon, Hoogduin, and Zhang, 2015). A false positive occurs when an audit procedure identifies an exception that later, through investigation, is determined to be a non-error that does not require additional follow up or any audit adjustment. In this study I investigate how the attributes of extent and precision influence auditors' judgments of errors which ultimately affect audit quality.

Auditors respond to the risk of material misstatement by adjusting the nature, timing, and extent of the audit procedures that they perform (PCAOB 2010b). The nature refers to which audit procedure to choose, the timing refers to when to perform the procedure, and the extent refers to the size and makeup of the sample. Due to cost and time constraints, auditors currently often choose an extent that is significantly less than 100 percent of the population, resulting in sampling (PCAOB 2017). The use of sampling relies on the assumption that "items in the population are…homogeneous in the sense that observation of some subset of items is useful for drawing conclusions about the

remainder of the population" (Burgstahler and Jiambalvo 1986, 233). Audit sampling assumes that errors are equally distributed throughout a population which allows auditors to "project the misstatement result of the sample to the items from which the sample was selected" (PCAOB 2017, AS 2315.26).[2]

Based on the theoretical assumption of equal error distribution, when a sample is randomly selected, the error amount projected from the results of the sample is expected to be equivalent to the true error amount in the entire population.[3] It may be assumed then that auditors' responses to errors will not differ when the extent of the procedure expands significantly, such as from a small sample to the entire population.[4] This is in line with prior psychology research findings that individuals often disregard sample sizes when they evaluate evidence (Kahneman and Tversky 1972;Tversky and Kahneman 1971). However, I expect evidence obtained from testing the entire population to be treated differently. Testing the entire population provides auditors with evidence that is more reliable for two reasons. First, evidence from a larger sample provides the auditor with greater coverage over the population, making it less likely that an error has gone

[2] This assumption holds for errors that are unintentional. Intentional errors are not expected to be randomly distributed.

[3] This is based on the theoretical assumption of equal error distribution. If errors are equally distributed, and a $100 error is found when testing 10% of the population, the projected error would be $1,000. If the entire population was tested, the assumption is that the error would also be $1,000, thus the two will be equal. While equal error distribution will not always hold in reality, auditors do not know the true error unless the entire population is tested, which is why they rely on error projection and the assumption of equal error distribution. Understanding how auditors would respond if the true error is later determined to be above or below the projected error is outside of the scope of the current research.

[4] The extent of the procedure could be expanded without testing the entire population. However, large extent is operationalized as testing the entire population because this study aims to contrast the two extreme conditions of a sample and the entire population.

undiscovered. Second, as a greater portion of the population is tested, less projection is necessary to extrapolate the discovered error to the remaining population resulting in an error that is more certain. Individuals have been found to be more influenced by concrete information (Kahneman and Tversky 1972) and auditors' negotiation is improved when their argument is more defensible (Magee and Tseng 1990). By providing auditors with a more defensible and reliable adjustment, testing a greater portion of the population is expected to increase the likelihood auditors will recommend adjusting the financial statements.

While expanding the extent of an audit procedure has benefits, it is also expected to impact the precision of the procedure by increasing auditors' exposure to false positives (Krahel and Titera 2015; Vasarhelyi, Kogan, and Tuttle 2015). False alarm literature, stemming from Aesop's Fable about the boy who cried wolf, suggests that repeated exposure to false alarms influences judgments of the warning system (Breznitz 1984). When a false alarm occurs, the warning system loses credibility, leading individuals to discount both the alarm and the potential danger. This is labeled the False Alarm Effect (FAE).

Returning to the audit context, the FAE suggests that auditors' judgments will be influenced by false positives in a way that can reduce audit quality. When the results of an audit procedure present auditors with frequent false positives, I expect that auditors will see the evidence provided by the procedure as less reliable. If an audit procedure provides unreliable evidence, then auditors' conclusions cannot rest on the results of the procedure. Additionally, exposure to numerous false positives is expected to cause auditors to discount true exceptions and fail to recognize the 'danger' of these exceptions

as they doubt the results of the procedure.[5] Together, the FAE will cause auditors to become skeptical of the audit procedure itself, and to rely less on the results of the procedure. The FAE is expected to be particularly strong when the extent is large, given that large extent audit procedures will expose auditors to a greater number of false positives, and repeated false alarms increase the strength of the FAE (Breznitz 1984). This aligns with the dilution effect which finds that the addition of nondiagnostic information reduces the magnitude of response to diagnostic information (Nisbett, Zukier, and Lemley 1981; Zukier 1982).

I administered a 2x2 experiment which manipulated Extent (sample vs. entire population) and False Positives (present vs. absent). Auditor participants were provided with the results of an audit procedure that included a summary of exceptions found. Participants learned the results of the investigation of exceptions which revealed whether each exception was an error or a non-error. Participants then made judgments including the likelihood of recommending that the client adjust the financial statements.

I predict and find that auditors are more likely to recommend audit adjustments for errors when using large extent procedures compared to small extent procedures. I find an interaction between Extent and False Positives, such that the effect of Extent on increasing the likelihood of recommending audit adjustments is *weaker* for procedures with false positives compared to procedures without false positives. Using PROCESS (Hayes 2017), I further find that the effect of Extent on likelihood judgments is mediated by auditors' perceptions of the reliability of the procedure, in addition to being moderated

[5] The term exception in this paper is used for an item that needs further investigation in order to determine if it is an error or not. The term error is used for an exception that was investigated and determined to be an error.

by the presence of false positives. Through a research question, I find that in my setting auditors are no more likely to recommend an audit adjustment when they use a large extent procedure that contains false positives compared to a small extent procedure with false positives. My findings suggest that false positives have the potential to diminish some of the benefits of large extent procedures.

This study contributes to academic theory as prior audit literature recognizes how "an important (and potentially scary) byproduct of frequent false alarms is that the output of a[n]… analysis may lose relevance for practicing auditors if it almost never uncovers an actual fraudulent entry" (Cleary and Thibodeau 2005, 80). Additionally, Issa and Kogan (2014) discuss how vast numbers of exceptions can overwhelm auditors and reduce audit efficiency and effectiveness. However, prior auditing literature has tested neither behavioral responses to false positives nor the ability of false positives to discredit procedures and reduce auditors' responsiveness to errors. The current study also contributes to the sampling literature (cf., Elder et al. 2013) and research on responses to errors projected from a sample (Burgstahler, Glover, and Jiambalvo 2000) by directly comparing auditors' responses to evidence from testing a sample and evidence from testing the entire population, while also incorporating the related variable of precision. Finally, this study contributes to the continuous auditing literature (e.g., Gonzalez and Hoffman 2018; Vasarhelyi and Halper 1991) as false positives can play a role in the continuous audit setting which aims to provide greater coverage through continuous monitoring.

The results are also relevant to audit practice because they identify how two attributes of audit procedures, extent and precision, alter auditors' treatments of audit

errors. Improving auditors' skepticism is a significant focus of audit research (cf., Hurtt et al. 2013), and a potential benefit of testing entire populations. In the absence of false positives, if auditors test the entire population and identify errors, auditors' skepticism is expected to increase as they place additional weight on errors found from the complete population audit procedure, compared to a sample audit procedure. However, my findings suggest that audit firms should take into consideration how false positives can negate the potential benefits of expanding the extent of audit procedures before they begin investing in Big Data analytics to test complete populations.

The rest of the paper is organized as follows. Section II outlines the background and hypothesis development, Section III explains the methodology, Section IV provides the results, and Section V concludes.

CHAPTER 2

BACKGROUND AND HYPOTHESIS DEVELOPMENT

Extent

Prior research has found that individuals often disregard sample size when evaluating evidence and instead make decisions based on the most salient characteristic regardless of the sample size (Tversky and Kahneman 1971; Kahneman and Tversky 1972). This suggests that an auditor will focus on the results of an audit procedure, such as the percentage of errors in the sample, when evaluating evidence and disregard the size of the sample used. However, sample sizes are a critical judgment made during the audit process in order to determine if auditors have gathered enough evidence, and the standards highlight the relevance of sample size to auditors' interpretation of the results of audit procedures (PCAOB 2017).

During the audit process, auditors obtain evidence to support the conclusion expressed in the auditor's report (PCAOB 2010b). In order for audit evidence to support the conclusion, the evidence must be both appropriate and sufficient (Messier, Glover, and Prawitt 2019). The *appropriateness* of audit evidence refers to the quality of the evidence; if evidence is not both relevant and reliable, it cannot provide support for the conclusion (PCAOB 2010b). In order to determine what quantity of audit evidence is considered *sufficient*, auditors must consider the risk of material misstatement, as well as the quality of the evidence (PCAOB 2010b). As the risk of material misstatement

increases, auditors need more evidence to support their conclusion. Auditors can increase or decrease the amount of evidence they gather in various ways including performing different tests and expanding the extent of their preexisting tests. When auditors choose to make extent changes to change the amount of evidence, discussions about the sufficiency of audit evidence directly affect the sample size decision (PCAOB 2017).

Auditors aim to select sample sizes that provide sufficient evidence in order to reduce sampling risk to the desired level. Sampling risk is the concern that a selected sample used will not be representative of the population, causing the auditor to draw an incorrect conclusion based on the sample (Messier, Glover, and Prawitt 2019). Sampling risk takes into consideration the reality that while equal error distribution is assumed on a theoretical level (Burgstahler and Jiambalvo 1986), it is not always true in practice. For example, a sample could indicate that the account is not materially misstated when it actually is (PCAOB 2017). Anecdotal evidence suggests auditors build concerns of sampling risk into their sample sizes. Increasing sample size allows auditors to reduce the likelihood that the auditor will fail to detect a material misstatement in the financial statements (i.e., detection risk).

Increasing the extent of an audit procedure increases the perceived quality of the evidence the audit procedure provides. This is due to the audit evidence from a larger sample being both more sufficient and more reliable to support the conclusion. Increases in extent make it less likely that the auditor will fail to uncover an error and more likely that the conclusion of the procedure is accurate. For this reason, expanding the extent is a tactic auditors' use to respond to risk, such as concerns of source credibility due to a client having low integrity (Beaulieu 2001). Therefore, in contrast to prior psychology

research (Kahneman and Tversky 1972; Tversky and Kahneman 1971), I expect auditors to consider the sample size when interpreting audit evidence.

Evidence Evaluation

When individuals evaluate evidence, the characteristics of the evidence influence the weight placed on it, such that information that is seen as more credible is more impactful during the decision process (Birnbaum, Wong, and Wong 1976). Users of financial statement information react to differences in reliability because accounting information must be reliable to be considered useful (Maines and Wahlen 2006). Auditors are sensitive to source reliability, and find information to be more diagnostic when it is from a more competent and independent source, such as a specialist (Hirst 1994). This discussion suggests that as the extent of the audit procedure affects perceptions of the reliability of the evidence, changes in extent will influence how auditors use and rely on the provided evidence.

Additionally, when auditors expand the sample size, this provides auditors with an error that is more certain because as extent increases, the amount of the error that must be projected to the untested population decreases. Having greater certainty over the error influences the strength of the evidence during discussions between auditors and clients because there is less ability for the client to dismiss a need for an adjustment. Individuals have been found to be more influenced by concrete information (Kahneman and Tversky 1972), and when there is less room for debate, auditors have greater power in negotiation (Magee and Tseng 1990). This is in line with prior literature's finding that auditors are more likely to waive adjustments when looking at subjective issues relating to estimates of future events, compared to objective issues (Braun 2001) and that whenever there is

greater uncertainty, it is expected that auditors will come to a conclusion that is more in line with the client's preference (e.g., Wright and Wright 1997).

Large extent audit procedures provide a more definitive misstatement and more error certainty. An error discovered through a large extent procedure is not easily disputed, whereas projected errors are less defensible. Thus, auditors will be more likely to recommend an audit adjustment with large extent procedures. I formally predict the following:

> H1: When an error is discovered, auditors will be more likely to recommend an audit adjustment when using a large extent audit procedure compared to a small extent audit procedure.

False Alarm Effect (FAE)

The goal of a warning system is to allow individuals to have knowledge of potential danger in hopes that they will have ample time to respond to the danger (Breznitz 1984). Warning systems exist in different capacities, from smoke detectors to anti-shoplifting devices placed on clothes. While warning systems are helpful, they are not 100 percent accurate, which can result in false alarms. In the boy who cried wolf fable, the boy cried "Wolf!" again and again, but when the villagers came to his rescue, no wolf was there. This caused the villagers to discount the boy's cries so that when he cried for help on the day when the wolf actually came, they failed to come to his rescue.

Repeated exposure to false alarms can influence how individuals respond to alarms (Breznitz 1984). For example, if a siren used to alert individuals of a nuclear power plant incident goes off accidentally multiple times, individuals may be less likely to respond to the siren the next time that they hear it (Mileti and Peek 2000). False alarms can alter individuals' judgments by causing individuals to discount the alarm which leads

them to ignore future alarms, and by causing individuals to discount the danger associated with the alarm (Breznitz 1984).

Audit procedures serve as a warning system to protect the auditor. The identification of an exception acts as an alarm to auditors that a material misstatement may exist in the financial statements. Not every audit procedure identifies an error, but when an error is found that is inconsistent with auditors' assumptions of risk, auditors are required to investigate the error and determine how it affects the audit plan (PCAOB 2017). This investigation aims to prevent auditors from providing a clean opinion on financial statements which contain a material misstatement.

In the audit context, false alarms are labeled false positives (Vasarhelyi, Kogan, and Tuttle 2015). These occur during the audit when auditors identify an exception that may be indicative of an error, but upon further investigation is determined to be correct. For example, during control testing, a procedure may uncover what are thought to be improper signoffs, but when the auditor investigates these exceptions, s/he learns that the manager responsible for the signoff was on vacation, and the 'improper' signoff was done by his assistant who was given signoff privileges while the manager is gone. Alternatively, during substantive testing, an auditor performing cutoff testing may identify an expense that s/he believes should have been accrued for, but is not in the accruals ledger. When the auditor requests additional supporting documentation for this exception, the auditor may learn that it is correct that the item was not accrued for because it pertains to the next fiscal year. In both of these false positive situations, the test operated as intended but the exception it identified had an explanation that indicated it

was not an error. When an exception is determined to be a false positive, there is no true reason for concern, and no further follow up or adjustments are necessary.

False positives suggest that the audit procedure is, at times, imprecise at identifying errors. Based on the FAE, I expect that exposure to false positives during testing will lead auditors to find the audit procedure to be less reliable compared to when they are not exposed to false positives. This is in line with prior literature's finding that auditors are skeptical of client provided support which contains small errors (Andiola et al. 2019). While false positives are not errors, I expect the exposure to false positives to reduce auditors' trust and reliance on the procedure, similar to when individuals process information that contains mistakes.

The presence of false positives and decrease in reliability of the procedure could lead auditors to become more skeptical, similar to when auditors face high fraud risk (Rose and Rose 2003) or due to auditors' inability to ignore invalidated error information (Ross, Lepper, and Hubbard 1975). Applying the FAE literature, the repeated false positives are instead anticipated to lead auditors to discount the audit procedure itself (Breznitz 1984). As auditors discount the audit procedure due to its unreliability, the conclusions of the audit procedure are expected to have less impact on their judgments. The presence of many false positives may lead auditors to discount the discovered errors, believing that even the errors are of little concern. This is similar to when individuals believe that a source is biased and discount the information gathered from that source by weighting it less (Birnbaum and Stegner 1979), and also aligns with the dilution effect. The dilution effect occurs when the addition of nondiagnostic information reduces the magnitude of response to the diagnostic information (Nisbett, Zukier, and Lemley 1981;

Zukier 1982). The dilution effect has been found to be powerful both outside of the accounting context (Liberman and Ross 2006; Tetlock and Boettger 1989) and within the accounting context (Fanning, Agoglia, and Piercey 2015; Glover 1997) and suggests that the presence of false positives will weaken auditors' responses to errors. Overall, as false positives make the audit procedure seem unreliable, auditors will reduce the weight they place on the evidence it provides.

While large extent procedures are hypothesized in H1 to increase the error response required by auditors, false positives are predicted to decrease error response. This leads to the following hypothesis;

> H2: When an error is discovered, auditors will be less likely to recommend an audit adjustment when using an audit procedure which contains false positives than an audit procedure which does not contain false positives.

While any exposure to false positives can affect individuals, the FAE is most likely and most powerful when individuals are exposed to repeated false alarms (Breznitz 1984). Assuming that false positives are equally distributed in the same way as errors (Burgstahler and Jiambalvo 1986), the frequency of auditors' exposure to false positives will increase when auditors increase the extent of a procedure to gain more coverage. Additionally, as auditors use computer assisted procedures to test the entire population, a failure to design a precise procedure will increase false positives (ICAEW 2016; Minniti and Camehl 2018; Vasarhelyi, Kogan, and Tuttle 2015). While H1 predicts that large extent procedures will increase the likelihood of recommending an audit adjustment, I expect false positives to mute this response. This is due to the large number of false positives in a large extent procedure reducing the perceived reliability of the procedure.

H3: When an error is discovered, the effect of extent on increasing the likelihood of recommending an audit adjustment will be weaker when false positives are present compared to absent.

Research Question

False positives are expected to mute auditors' error response, but prior literature does not provide a basis on which to predict how much false positives will alter the value of large extent audit procedures. Regardless of the presence or absence of false positives, the expanded extent will provide auditors with a more concrete error, compared to an audit procedure that tests a smaller sample. Due to this, auditors may be more likely to recommend an adjustment for a large extent audit procedure compared to a small extent audit procedure, regardless of the presence of false positives. Comparing small extent procedures with false positives to large extent procedures with false positives is interesting as it compares the current normal protocol to the proposed future protocol. This will provide information regarding the potential benefit of expanding sample sizes, given the likelihood that false positives will consistently be present no matter the chosen extent. As there is no theoretical background for comparing the effects of false positives and extent, I propose the following research question:

RQ: When an error is discovered, will auditors be more likely to recommend an audit adjustment when using large extent audit procedures with false positives compared to small extent audit procedures with false positives?

CHAPTER 3

METHOD

Participants

The study was administered to auditors working at a regional public accounting firm. In total, 113 participants completed the study using Qualtrics during a required firm training.[6] Participant experience ranged from staff to manager auditors, with an average of two years of full time audit experience. 55.6% of participants are males and 44.4% are females.[7]

Independent Variables

The experiment utilized a 2x2 between-participant experimental design that manipulated the Extent of the audit procedure (sample vs. entire population) and False Positives (present vs. absent). In the Sample conditions, the audit procedure tested 20% of all revenue transactions, and in the Population conditions, the audit procedure tested the entire population of revenue transactions.[8] In the False Positives conditions, the initial results of the audit procedure included false positives, while in the No False Positives conditions, the initial results of the audit procedure did not include false positives.

[6] Two participants began the survey but did not finish, so they were not included in the analyses due to incomplete data.
[7] One participant chose not to provide gender information.
[8] By design, by testing 20% of the individual sales, the auditor tests 20% of the dollar value of revenue.

In all conditions, 300 revenue transactions occurred at the company, Seven Seas, which designs, manufactures, and sells premium yachts. Unknown to the participants, 5% of sales transactions were recorded with errors. In all conditions, auditors were provided with information that contained the initial results of the audit procedure. The initial results informed participants about the number of exceptions found, but participants did not learn how many exceptions were errors or non-errors until later.[9, 10] The content of the initial results provided to the participants varied depending on the randomly assigned condition.

In the Sampling/No False Positives condition, 60 items were tested and three errors were discovered. In the Population/No False Positives condition, 300 items were tested and 15 errors were found. In the False Positives conditions, there was an additional 15% false positives rate beyond the 5% error rate.[11] Therefore, in the Sample/False Positives condition, there were nine false positives found, and in the Population/False Positives condition, there were 45 false positives found. These false positives were all later determined to be 'non-errors'. For False Positives conditions, the total number of exceptions in the initial results was the sum of the errors and false positives. For No False

[9] In practice, auditors are required to investigate all exceptions identified and for any exception that is determined to be a non-error, and thus a false positive, document the reason that it is a non-error. Therefore, it is realistic that the initial results of an audit procedure will include both errors and false positives.

[10] In order to avoid any negative connotations tied to the phrase 'false positive', the term non-error is used in the instrument rather than false positive.

[11] Part of the motivation of this study is the concern that expanding extent will lead to greater numbers of false positives than previously experienced. Pilot data revealed that participants anticipate a high false positive rate of approximately 50%. The false positive rate in the study is set slightly above this to test the effects of high false positive rates. With 5% errors and 15% false positives, participants in the false positive conditions learn that 25% of the exceptions are errors and 75% are false positives.

Positives conditions, the total number of exceptions in the initial results was the sum of only the errors.

Task and Procedures

The materials used in this experiment were based in part on Burgstahler, Glover, and Jiambalvo (2000) and the related case (Beasley et al. 2019). The materials were pilot tested with audit participants, adjusted based on the findings of the pilot test, and reviewed by a senior audit associate. Feedback related to the instrument design and appropriateness of the task for the experience level of participants was also provided by a senior audit associate, audit manager, and senior manager.

First, participants read background information about the company, Seven Seas. The materials informed participants that they were interpreting the results of a substantive test performed over the Revenue account, specifically a three-way match agreeing the customer order, shipping document, and billing document. They were told that the company did not have any weaknesses in internal controls. They were provided with the materiality level, which was set at 5% of the company's pretax income, and tolerable misstatement which was set at 50% of materiality (Eilifsen and Messier 2015). To deter participants from interpreting different risk levels based on the extent of the procedure, they were told that the initial estimate, prior to testing, of the likelihood of Revenue being misstated was 20%, in line with Andiola et al. (2019).

After reading the introductory information, participants were provided with a document that summarized the initial results of the substantive procedure performed. In all conditions, participants were told how many exceptions were found, the client balance of those accounts, the audited balance of those accounts, and the difference between the

client balance and the audited balance. In the Sample conditions, the difference was projected to the population, and the projected error equaled the difference found in the Population conditions.[12] The projection amounts stayed silent on sampling risk to avoid incorporating additional differences between conditions and due to the fact that the participants' firm incorporates sampling risk into the sample size, not into the extrapolation of the errors. The pre-investigation difference was significantly above tolerable misstatement for the False Positives conditions, but only slightly above tolerable misstatement for the No False Positives conditions. In all conditions the pre-investigation difference was below overall materiality.

Participants were then informed that each exception had been investigated to determine if it was an error, and were provided with an explanation of what an error and a non-error were prior to reviewing the results.[13] For each exception, participants were presented with a screen that had information about the specific exception, including the customer name, client balance, audited balance, and difference. Then on the next screen the participants were told whether the exception was an error or non-error. In order to facilitate the speed of reviewing the exceptions, participants were told to focus on the results of the investigation rather than the numeric amounts, and colors were used to re-enforce the finding (i.e., green for non-error and red for error). The number of exceptions reviewed corresponded with the participant's respective condition, only the False

[12] The Sample/False Positives projection equals the Population/False Positives amount and the Sample/No False Positives amount equals the Population/No False Positives projection. Inherent in the design, the False Positives conditions have higher initial exception amounts.

[13] Participants were not provided with an explanation for why the non-error occurred due to the fact that who was at fault, auditor or client, could interact with false positives and that is outside of the scope of this study.

Positives conditions included non-errors, and the exception order was randomized for each participant. After reviewing all of the exceptions, participants were provided with a summary screen which included the error exception amounts, non-error exception amounts, and non-exception amounts.[14]

In the summary screen, for Sample conditions the error amount was projected to the population, and once projected, equaled the discovered difference in the Population conditions. The post-investigation difference was slightly above tolerable misstatement, below overall materiality, and was equal in all conditions. In the No False Positives conditions, the post-investigation difference was equal to the initial difference, and in the False Positives conditions, the post-investigation difference was lower than the initial difference. Please see Appendix A for excerpts from the instrument.

Dependent Variables

Once participants reviewed the final summary information, they responded to the dependent variable questions. First, participants were asked the likelihood they would recommend the client to adjust the Revenue account prior to issuing the financial statements on a 0 – 100 slider ranging from "Very Unlikely" to "Very Likely" and how much of an adjustment they would recommend.[15] Auditors were next asked how likely they believed it is that Revenue is materially misstated on a 10-point Likert scale ranging from "Very Unlikely" to "Very Likely." Participants indicated the likelihood they would

[14] Non-exception amounts are items that were tested in the audit procedure and did not have exceptions, and therefore were not part of the exceptions identified.
[15] In the Population condition, auditors are provided with a known error; however prior research has found that even when an error amount is known, there is variation in auditors' adjustment amounts (Libby and Brown 2013; Kachelmeier and Van Landuyt 2017). Therefore, this question is relevant in all conditions.

perform additional substantive procedures, and if so, what additional substantive procedures they would perform. Auditor participants then selected a response on a scale from 0 "Not Reliable" to 10 "Very Reliable" to the question, "How reliable do you believe the audit procedure results are?" as well as answered how defensible they believe their proposed adjustment was to the client.

After completing all dependent variable questions, participants responded to manipulation check questions. The first manipulation check question asked participants to indicate whether the audit procedure was performed over the entire population or a 20% sample. The second manipulation check question asked participants if there were non-errors included in the audit procedure results. Participants also responded to a true/false question asking if the procedure performed over Revenue was a three-way match. Following manipulation checks, participants responded to post experimental questions including the Rotter Interpersonal Trust Scale (RIT) (Rotter 1967) and demographic questions, including their level and the number of clients they serve each year.[16] The materials for participants in the Population – False Positives condition were significantly longer than in the other conditions. To avoid Population – False Positives condition participants from noting the time difference, all other participants received an additional block of questions after they completed the demographic questions. This block had slider questions unrelated to the task they performed and was collected after everything else to avoid it interacting with any of their answers.

[16] The RIT was chosen over the Hurtt Professional Skepticism Scale (Hurtt 2010) as Quadackers, Groot, and Wright (2014) finds that inversed RIT predicts auditors' skeptical judgments and actions better than the Hurtt scale, over a variety of risk situations.

CHAPTER 4

RESULTS

Manipulation Checks

Of the 113 participants, 17 participants (15%) failed the extent manipulation check, 11 (10%) failed the presence of false positives manipulation check, and five (4%) answered the true/false question regarding the type of procedure performed incorrectly. Of these, five participants (4%) answered two questions wrong, but no participant missed all three questions. Results are robust to removing participants who missed one or two manipulation check questions. Therefore, I did not remove these participants.

Hypothesis Tests H1 – H3

To test my hypotheses, I ran an ANCOVA with *Extent* and *False Positives* as the independent variables, *Likelihood* as the dependent variable, and three covariates (Inverse RIT, Level, and Client Number).[17] *Likelihood* captures participants' likelihood of recommending that the client adjusts the Revenue account prior to issuing the financial statements. RIT was inversed in line with prior research (Quadackers, Groot, and Wright 2014), Level is based on the title of their current position, and Client Number captures the number of clients they serve each year. The results of the ANCOVA are tabulated in Table 4.1.

[17] An ANCOVA was run with multiple possible covariates, including those above, in order to determine if any were significant. The three above (i.e., Inverse RIT, Client Number, and Level) are significant, so they are maintained for the final model. No other covariates were found to be significant.

H1 predicts that auditors will be more likely to recommend an audit adjustment when an error is found using a large extent audit procedure compared to a small extent audit procedure. Panel A of Table 4.1 presents the descriptive statistics from the ANCOVA. Consistent with H1, the mean *Likelihood* for the Population and Sample conditions is 88.13 and 74.57, respectively. Panel B of Table 4.1 presents the ANCOVA results, which indicate that *Extent* has a significant effect on *Likelihood* ($F = 9.75$; $p = 0.001$, one-tailed). This finding provides strong support for H1. Since an interaction is predicted in H3, this main effect finding should be interpreted carefully in light of the anticipated interaction.

H2 predicts auditors will be less likely to recommend an audit adjustment when an error is discovered using an audit procedure which contains false positives compared to an audit procedure which does not contain false positives. As shown in Panel A of Table 4.1, and inconsistent with H2, the mean *Likelihood* for the False Positives Absent and False Positives Present conditions is 79.79 and 82.57, respectively. The ANCOVA results in Panel B of Table 4.1 indicate that *False Positives* does not have a significant effect on *Likelihood* ($F = 0.01$; $p = 0.543$, one-tailed), indicating that the presence of false positives alone does not significantly influence likelihood judgments.[18]

H3 predicts that the effect of *Extent* on increasing *Likelihood* will be weaker when *False Positives* are present compared to absent. Consistent with H3, Figure 4.1 illustrates a significant interaction between *Extent* and *False Positives*, and this is supported by the ANCOVA results which indicate a significant interaction ($F = 4.35$; $p = 0.020$, one-

[18] This p-value has been adjusted to reflect the fact that the results are in the opposite direction from what was predicted. Calculated as 1 – (0.914/2).

tailed). Follow-up tests of simple effects found in Panel C of Table 4.1 show a significant simple effect of *Extent* when *False Positives* are absent ($F = 13.66$, $p < .001$, one-tailed) but not when *False Positives* are present ($F = 0.56$, $p = 0.456$, two-tailed). These results provide strong support for H3.

In order to gain additional confidence in these findings, I ran the Kruskal-Wallis Test, a rank-based non-parametric test, to determine if there were differences in median Likelihood scores between the four conditions. Median scores were statistically different between groups, ($\chi^2(3) = 13.783$, $p = 0.003$, two-tailed, untabulated), which indicates that at least two groups differ in their medians. Select pairwise comparisons were performed using Dunn's (1964) procedure. This post hoc analysis identified a statistically significant difference in median Likelihood scores between Population – No False Positives and Sample – No False Positives conditions ($p < 0.001$, one-tailed, untabulated). It did not reveal statistically significant differences between Population – False Positives and Sample – False Positives conditions ($p = 0.109$, two-tailed, untabulated).[19] This provides further support that when false positives are absent, individuals are more likely to recommend an adjustment. In the presence of false positives, this difference is not found.[20]

[19] This p-value cannot be divided by two as it pertains to the research question where no directional prediction was made.

[20] As a third test of H3, I used regression analysis. H3 is statistically equivalent to the expectation that size of the regression coefficient for Extent should be larger when False Positives are absent compared to present. In order to test this, I first ran two regressions, one with the data for False Positives absent only and one with the data for False Positives present only. When False Positives are absent, the coefficient for Extent is 24.860 ($p < 0.001$, one-tailed, untabulated) whereas when False Positives are present, the coefficient for Extent is 5.569 ($p = 0.392$, two-tailed, untabulated). The results appear to suggest that Extent is a much stronger predictor of Likelihood when False Positives are absent compared to present. To formally test this, I ran a regression with the complete set of data

Pairwise comparison results can be found in Table 4.2. Comparing between cells for the Population condition, the *Likelihood* marginally decreases when false positives are introduced (p = 0.086, one-tailed). For the Sample condition, the difference is in the opposite direction as predicted with Sample – No False Positives having a greater *Likelihood* than Sample – False Positives, however this difference is not significant (p = 0.941, one-tailed).[21] The *Likelihood* is greater for Population – False Positives than Sample – No False Positives, (p = 0.021, two-tailed), which suggests that there can still be a benefit to testing the entire population, even when false positives are present.

Moderated Mediation

H1 finds that as extent increases, auditors are more likely to recommend an adjustment. H3 further finds that this effect is moderated by the presence or absence of false positives. In order to better understand why and how this occurs, I tested for moderated mediation using PROCESS (Hayes 2017) with participants' responses to the question "How reliable do you believe the audit procedure results are?", labeled as *Reliability*, serving as the mediator.

Based on the theoretical development for H1 – H3, Model 8 best fits the results that I anticipated because it expects both Extent and False Positives to influence the perceived Reliability. Model 8 predicts that *Extent*, *False Positives*, and the *Extent X*

---

with Extent, False Positives, and Extent*False Positives as independent variables. In this regression, the interaction coefficient is -19.28 (p = 0.020, one-tailed, untabulated). The significance of this coefficient indicates that the regression coefficient for False Positives present is significantly different than for False Positives absent. This pattern of results further supports H3 and that Extent will have a greater effect on increasing Likelihood when False Positives are absent compared to present. All regressions included the three covariates for consistency.

[21] This p-value has been adjusted to reflect the fact that the results are in the opposite direction from what was predicted. Calculated as 1 – (.118/2).

False Positives interaction will each influence both *Reliability* and *Likelihood*. The covariates from the original ANCOVA were included for consistency. The conceptual diagram is depicted in Figure 4.2.

The results for Model 8 of PROCESS, run with 5,000 bootstrapped samples, are found in Table 4.3.[22] Panel A shows that *Extent* significantly influences *Reliability* ($p < 0.001$, one-tailed), however, *False Positives* and *Extent X False Positives* do not ($p = 0.868$, one-tailed and 0.165, one-tailed, respectively).[23] As shown in Panel B, *Reliability* serves as a mediator as it significantly affects *Likelihood* judgments ($p = 0.001$, one-tailed). *False Positives* do not significantly affect *Likelihood* ($p = 0.672$, one-tailed), however the interaction does ($p = 0.032$, one-tailed).[24] As *Extent* significantly influences *Likelihood* ($p = 0.028$, one-tailed) even when *Reliability* is included, these results indicate *Reliability* partially mediates the relationship between *Extent* and *Likelihood*. Panel C of Table 4.3 shows that, in line with the interaction found in H3, the effect of *Extent* does vary depending on the level of *False Positives*, which combined with the mediator of *Reliability* found above, suggests moderated mediation. However, in Panel D of Table 4.3, the confidence interval testing for moderated mediation contains 0 which indicates that the results do not provide evidence for the type of moderated mediation expected of Model 8.

[22] In line with Chapter 8 of Hayes (2017), Extent and False Positive Presence were coded as -.5 and .5 rather than dummy coded as 0 and 1. This allows the results of PROCESS to be comparable to a 2x2 ANCOVA.
[23] The *False Positives* p-value has been adjusted due to the results being in the opposite direction as predicted. Calculated as 1 – (.264/2)
[24] The *False Positives* p-value has been adjusted due to the results being in the opposite direction as predicted. Calculated as 1 – (.656/2)

While the theory suggested that false positives would influence perceived reliability, the results above indicate that neither *False Positives* nor the *Extent X False Positives* interaction to influence *Reliability*. This pattern of results is consistent with Model 5 which is identical to Model 8 except removing these two relations. Testing Model 5 allows me to provide support for moderated mediation, but of a different type. The conceptual diagram is depicted in Figure 4.3. The results of testing this model with 5,000 bootstrapped samples are presented in Table 4.4.

Panel A of Table 4.4 shows that, in line with the previous results, *Extent* significantly influences *Reliability* ($p < 0.001$, one-tailed). The results in Panel B of Table 4.4 are identical to those in Model 8 such that *Reliability* ($p = 0.001$, one-tailed), *Extent* ($p = 0.028$, one-tailed), and *Extent X False Positives* ($p = 0.032$, one-tailed) significantly influence *Likelihood*, but *False Positives* do not ($p = 0.672$, one-tailed).[25] Panel C provides support for the moderation as, in line with the simple effects found when testing H3, *Extent* is only significant when *False Positives* are absent ($p = 0.005$, one-tailed), but not when they are present ($p = 0.881$, two-tailed). Panel D provides support for the indirect effect of *Extent* on *Likelihood* through *Reliability*. Together, Panel C and D provide support for moderated mediation, in line with Hayes (2017) Model 5. Specifically, *Reliability* mediates the relationship between *Extent* and *Likelihood*, and the effect of *Extent* on *Likelihood* is moderated by *False Positives*.

[25] The *False Positives* p-value has been adjusted due to the results being in the opposite direction as predicted. Calculated as 1 – (.656/2)

Research Question

The research question investigates whether when an error is discovered, auditors will be more likely to recommend an audit adjustment when using large extent audit procedures with false positives compared to small extent audit procedures with false positives. I compare the mean *Likelihoods* for Population – False Positives and Sample – False Positives, which are 85.04 and 80.11, respectively. As shown in Table 4.2, in these two conditions when false positives are present, the difference is not significant ($p = 0.456$, two-tailed). This shows that when there are false positives, the likelihood for recommending an adjustment does not differ between the Population and Sample Conditions. In contrast, when there are no false positives, the *Likelihood* is greater for the Population condition compared to the Sample condition ($p < 0.001$, one-tailed). This provides evidence that in my setting, the benefit of testing the entire population is diminished when false positives are present. Figure 4.1 illustrates this relationship graphically.

Supplemental Analyses

In addition to capturing the likelihood of recommending an adjustment, participants indicated how likely they would be to perform additional substantive procedures. I ran a 2x2 ANCOVA with *Additional Procedures* as the dependent variable, *Extent* and *False Positives* as the independent variables, and the three covariates (Inverse RIT, Level, and Client Number). The mean *Additional Procedures* is 7.95 for the Sample Condition and 6.67 for the Population Condition. The ANCOVA reveals a significant effect of *Extent* ($F = 4.71$, $p = 0.016$, one-tailed, untabulated), such that when *Extent* is large, participants are *less* likely to perform additional procedures. As a common

additional procedure would be to select additional samples to test, this aligns with expectations due to the fact that large extent procedures already tested the entire population. While directionally the results of *False Positives* matched expectations, such that when they are present the mean is higher than when they are absent (7.45 and 7.21, respectively), there is no significant effect of *False Positives* ($F = 0.02$, $p = 0.449$, one-tailed, untabulated). It is possible that participants in the False Positives conditions did not plan to perform more additional substantive procedures because they felt the additional procedures would also contain false positives and would not aid in providing the additional evidence desired. The interaction is also not significant ($F = 0.005$, $p = 0.943$, two-tailed untabulated).

Participants also provided a dollar value for their recommended adjustment. I ran a 2x2 ANCOVA with *Dollar Adjustment* as the dependent variable, *Extent* and *False Positives* as the independent variables, and the three covariates. The average *Dollar Adjustment* is $705,788 in the Sample Condition and $1,084,577 in the Population Condition, which is in line with expectations that the adjustment would be higher when the entire population is tested. The ANCOVA supports a significant effect of *Extent* ($F = 13.52$, $p < 0.001$, one-tailed, untabulated). The average *Dollar Adjustment* is $844,942 when false positives are absent and $937,693 when false positives are present, which is not in line with expectations that adjustments would be lower when false positives are present due to discounting the information. This pattern of results may suggest that participants are unable to ignore invalidated error information (Ross, Lepper, and Hubbard 1975), leading them to propose directionally higher adjustments. While directionally higher in the False Positives conditions, the effect of *False Positives* is not

significant (F = 0.42, p = 0.740, one-tailed, untabulated).[26] Additionally, the interaction (F = 0.01, p = 0.930, two-tailed, untabulated) is not significant.[27] The results indicate that the dollar value adjustment proposed would be higher when the entire population is tested, compared to a sample.

I also ran an ANCOVA identical to above, but with participants' judgments about the likelihood that the account is materially misstated, *Materially Misstated*, as the dependent variable and found that *Extent* is significant (F = 2.89, p = 0.046, one-tailed, untabulated), while *False Positives* (F = 0.35, p = 0.851, two-tailed, untabulated) and the interaction (F = 0.01, p = 0.912, two-tailed, untabulated) are not. The results indicate participants felt it is more likely the account is materially misstated when the entire population is tested, compared to a sample.

[26] The *False Positives* p-value has been adjusted due to the results being in the opposite direction as predicted. Calculated as 1 – (.521/2)

[27] 2 outliers were removed from the sample when this analysis was performed because their responses were more than 10 standard deviations away from the mean. Maintaining these participants in the sample does change the results such that Extent is no longer significant (F = 1.04, p = 0.155, one-tailed). These participants provided adjustment amounts of $6,656,125 and $18,000,000 which are 5 and over 13 times the discovered/projected misstatement of $1,331,225. Removing these participants from the ANCOVA used to test H1- H3 does not affect the conclusions.

Table 4.1 ANCOVA: Main Effects and Interactions

Panel A: Mean (std. dev.) [n] by condition

	False Positives Absent	False Positives Present	Row Means
Extent - Population	91.33 (16.71) [27]	85.04 (26.75) [28]	88.13 (22.41) [55]
Extent - Sample	69.40 (32.45) [30]	80.11 (20.94) [28]	74.57 (27.80) [58]
Column Means	79.79 (28.23) [57]	82.57 (23.93) [56]	81.17 (26.11) [113]

Panel B: Analysis of variance

Source	df	M.S.	*F*-statistic	*p*-value
Extent	1	5,789.85	9.75	**0.001**
FP Presence	1	7.02	0.01	**0.543**
FP*Extent	1	2,583.44	4.35	**0.020**
Covariates:				
Inverse RIT	1	1,866.45	3.14	0.079
Level	1	2,959.31	4.99	0.028
Client Number	1	1,889.25	3.18	0.077
Error	106	593.70		

Panel C: Simple effect of Extent given presence or absence of false positives

Effect of Extent given:	*F*-statistic	*p*-Value
False Positives Absent	13.66	**<.001**
False Positives Present	0.56	0.456

Table 4.1 reports the results of the measure of likelihood of recommending an audit adjustment. Reported *p*-values for directional predictions are one-tailed. This is indicated using **boldface.** Likelihood of Recommending an Audit Adjustment was captured as the answer to how likely they would be to recommend an audit adjustment prior to issuing the financial statements, captured on a scale of 0 – 100 ranging from "Very Unlikely" to "Very Likely". Inverse RIT is the inverse of the sum of the RIT questions. Level is the participant's current position, ranging in the sample from Intern to Manager, translated to a number ranging from 1-4. Client Number is approximately how many clients they serve each year. The means in Panel A have not been adjusted for covariates. The p-value for FP Presence has been adjusted to reflect the fact that the results are in the opposite direction from what was predicted. Calculated as 1 – (0.914/2).

Extent = Extent (Sample or Population)
FP = False Positives (Present or Absent).

Table 4.2 Pairwise Comparisons

Comparison	Difference [a]	*p*-Value
PNFP vs. SNFP	24.21	**<.001**
PFP vs. SFP (RQ)	4.93	0.456
PNFP vs. PFP	9.14	**0.086**
SNFP vs. SFP	-10.14	**0.941**
PFP vs. SNFP	15.08	0.021
PNFP vs. SFP	14.07	**0.019**

[a] These differences are based on mean amounts which have covariates evaluated at the following values: Inverse RIT = 0.0135, Level = 2.43, Client Number = 16.63.

Reported *p*-values for directional predictions are one-tailed. This is indicated using **boldface.** The p-value for SNFP vs. SFP has been adjusted to reflect the fact that the results are in the opposite direction from what was predicted. Calculated as 1 – (.118/2).

Extent = Extent (Sample or Population)
FP = False Positives (Present or Absent).
RQ = Research Question.

Conditions Key:
PNFP: Population – No False Positives
SNFP: Sample – No False Positives
PFP: Population – False Positives
SFP: Sample – False Positives

Table 4.3 Moderated Mediation – Model 8

Panel A: Outcome Variable - Reliability

	Coefficient	t	p-value
Constant	6.24	2.90	0.005
Extent	1.05	3.91	< **0.001**
FP	-0.30	-1.12	**0.868**
Extent*FP	-0.52	-0.98	**0.165**
Covariates:			
Inverse RIT	19.10	0.12	0.904
Level	0.22	1.05	0.297
Client Number	0.04	3.26	0.002

Panel B: Outcome Variable – Likelihood

	Coefficient	t	p-value
Constant	-1.43	-0.04	0.970
Extent	9.30	1.94	**0.028**
FP	2.00	0.45	**0.672**
Extent*FP	-16.67	-1.87	**0.032**
Reliability	5.03	3.09	**0.001**
Covariates:			
Inverse RIT	4,779.57	1.81	0.074
Level	-9.24	-2.62	0.010
Client Number	0.16	0.84	0.405

Panel C: Conditional effects of Extent given presence or absence of false positives

Effect of Extent given:	Effect	*p*-Value
False Positives Absent	17.64	**0.005**
False Positives Present	0.97	0.881

Panel D: Index of Moderated Mediation

	Index	BootSE	BootLLCI	BootULCI
False Positives	-2.61	3.02	-9.13	2.79

Reported *p*-values for directional predictions are one-tailed. This is indicated using **boldface.** Likelihood was captured as the answer to "How likely they would be to recommend an audit adjustment prior to issuing the financial statements?" on a scale of 0 – 100 ranging from "Very Unlikely" to "Very Likely." Inverse RIT is the inverse of the sum of the RIT questions. Level is the participant's current position, ranging in the sample from Intern to Manager, translated to a number ranging from 1-4. Client Number is approximately how many clients they serve each year. In Panel A, the False Positives p-value has been adjusted due to the results being in the opposite direction from what was predicted. Calculated as 1 – (.264/2). In Panel B, the False Positives p-value has been adjusted due to the results being in the opposite direction from what was predicted. Calculated as 1 – (.656/2).

Extent = Extent (Sample or Population)
FP = False Positives (Present or Absent)

Table 4.4 Moderated Mediation – Model 5

Panel A: Outcome Variable - Reliability

	Coefficient	t	p-value
Constant	6.48	3.02	0.003
Extent	1.04	3.88	**< 0.001**
Covariates:			
Inverse RIT	0.04	< .01	0.999
Level	0.24	1.15	0.254
Client Number	0.03	3.12	0.002

Panel B: Outcome Variable – Likelihood

	Coefficient	t	p-value
Constant	-1.43	-0.04	0.970
Extent	9.30	1.94	**0.028**
FP	2.00	0.45	**0.672**
Extent*FP	-16.67	-1.87	**0.032**
Reliability	5.03	3.09	**0.001**
Covariates:			
Inverse RIT	4,779.57	1.81	0.074
Level	-9.24	-2.62	0.010
Client Number	0.16	0.84	0.405

Panel C: Conditional effects of Extent given presence or absence of false positives

Effect of Extent given:	Effect	*p*-Value
False Positives Absent	17.64	**0.005**
False Positives Present	0.97	0.881

Panel D: Indirect Effect of Extent on Likelihood

	Effect	BootSE	BootLLCI	BootULCI
Reliability	5.23	2.36	0.96	10.28

Reported *p*-values for directional predictions are one-tailed. This is indicated using **boldface.** Reliability was captured as the answer to "How reliable do you believe the audit procedure results are?" on a scale from 0 – 10 ranging from "Not Reliable" to "Very Reliable." Likelihood was captured as the answer to "How likely they would be to recommend an audit adjustment prior to issuing the financial statements?" on a scale of 0 – 100 ranging from "Very Unlikely" to "Very Likely." Inverse RIT is the inverse of the sum of the RIT questions. Level is the participant's current position, ranging in the sample from Intern to Manager, translated to a number ranging from 1-4. Client Number is approximately how many clients they serve each year. In Panel B, The False Positives p-value has been adjusted due to the results being in the opposite direction from what was predicted. Calculated as 1 – (.656/2)

Extent = Extent (Sample or Population)
FP = False Positives (Present or Absent)

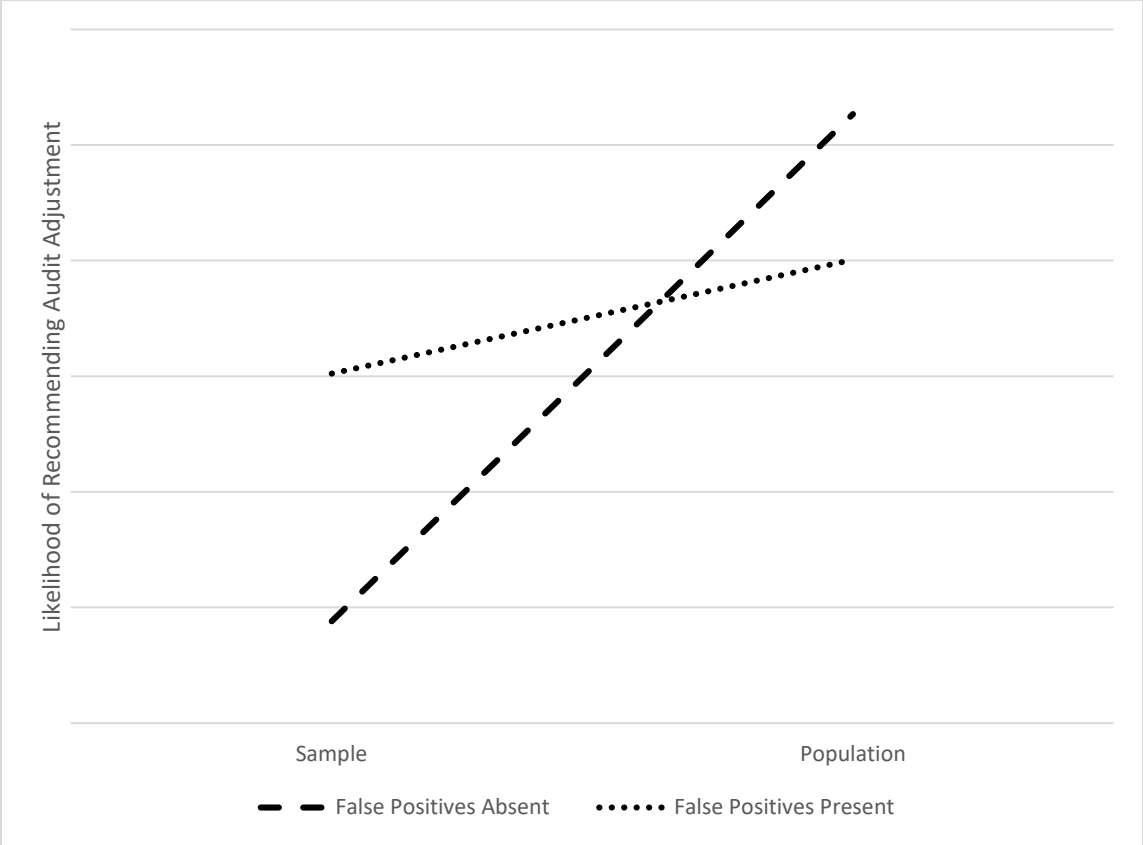

Figure 4.1: Graph of Results

Figure 4.1 graphs the likelihood that the auditor would recommend an adjustment prior to issuing the financial statements. Likelihood of Recommending an Audit Adjustment was captured as the answer to how likely they would be to recommend an audit adjustment prior to issuing the financial statements, captured on a scale of 0 – 100 ranging from "Very Unlikely" to "Very Likely."

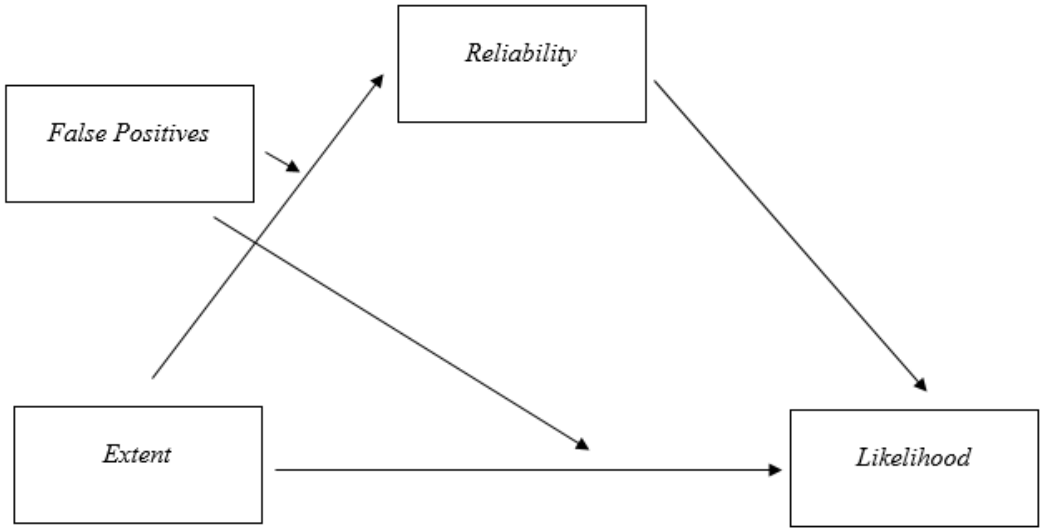

Figure 4.2: Moderated Mediation – Model 8

Figure 4.2 illustrates the conceptual diagram of Model 8 (Hayes 2017). *Reliability* was captured as the answer to "How reliable do you believe the audit procedure results are?" on a scale from 0 – 10 ranging from "Not Reliable" to "Very Reliable." *Likelihood* was captured as the answer to "How likely they would be to recommend an audit adjustment prior to issuing the financial statements?" on a scale of 0 – 100 ranging from "Very Unlikely" to "Very Likely."

Extent = Extent (Sample or Population)
False Positives = False Positives (Present or Absent)

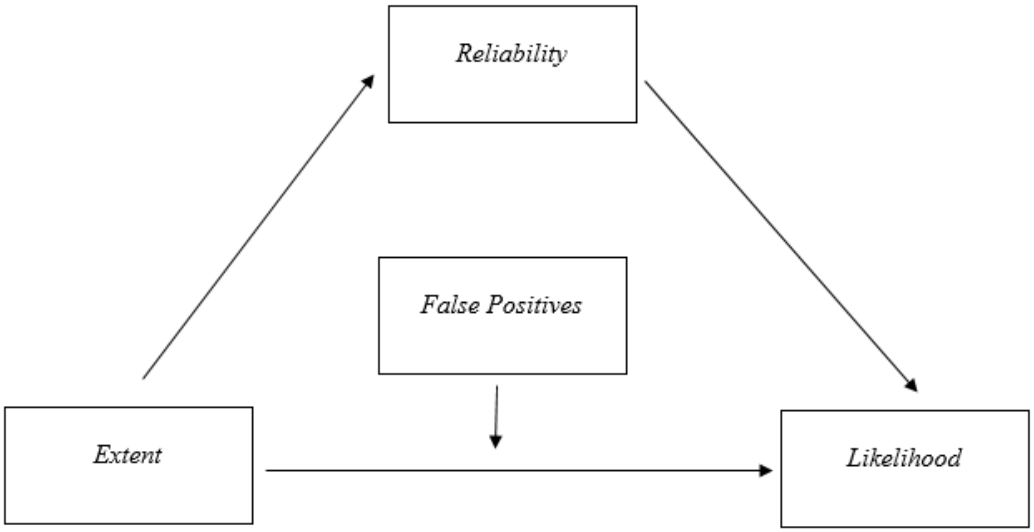

Figure 4.3: Moderated Mediation – Model 5

Figure 4.3 illustrates the conceptual diagram of Model 5 (Hayes 2017). *Reliability* was captured as the answer to "How reliable do you believe the audit procedure results are?" on a scale from 0 – 10 ranging from "Not Reliable" to "Very Reliable*." Likelihood* was captured as the answer to "How likely they would be to recommend an audit adjustment prior to issuing the financial statements?" on a scale of 0 – 100 ranging from "Very Unlikely" to "Very Likely."

Extent = Extent (Sample or Population)
False Positives = False Positives (Present or Absent)

CHAPTER 5

CONCLUSION

The results of this study aim to provide insight into how two attributes of the audit procedure, extent and precision, influence auditors' judgments and decisions. By operationalizing precision through the presence or absence of false positives, this paper speaks to concerns over the use of Big Data analytics to test entire populations leading to large numbers of false positive exceptions (Cao, Chychyla, and Stewart 2015; Krahel and Titera 2015; Minniti and Camehl 2018; Vasarhelyi, Kogan, and Tuttle 2015). I find that Extent influences how likely it is that auditors will recommend the client to record an adjustment, with the likelihood being higher when extent is large compared to small. This is moderated by False Positives such that when false positives are absent, Extent influences auditors' judgments; however, when false positives are present, auditors are no more likely to recommend an adjustment when extent is large compared to small. The effect of Extent on likelihood judgements is mediated by reliability perceptions. The results suggest that false positives have the potential to reduce the benefits of increasing the extent of audit procedures.

Auditor error response is one of the most important parts of the audit process. If auditors discover errors but do not respond appropriately, they fail to satisfy their responsibilities. The ability to use Big Data is predicted to aid auditors in discovering

errors and to improve audit quality (Yoon, Hoogduin, and Zhang 2015). One of the greatest benefits of Big Data analytics is that it will allow auditors to test entire populations rather than a sample; one of the greatest risks is that due to imprecision, it may lead to the discovery of numerous false positives (Cao, Chychyla, and Stewart 2015; Minniti and Camehl 2018) which based on my findings, could hinder reliance on Big Data analytics audit procedures.

My findings contribute to and extend theory. FAE literature has primarily been used in testing alert systems that serve to prevent catastrophic events, such as air traffic conflicts (Wickens et al. 2009) or nuclear disasters (O'Hara and Brown 1991). The current paper extends FAE into the audit setting, providing evidence that in non-life threatening settings, the FAE is strong enough to influence judgments. Using FAE, this study tests, and aims to support, the concern that was raised by Cleary and Thibodeau (2005) that frequent false alarms may cause audit tests to lose significance. It also contributes to both sampling literature and continuous auditing literature by identifying the role that false positives can play in hindering the benefits of extended sample sizes.

The results of this study are relevant for practitioners and regulators so they can understand both the benefits and risks of expanding the extent of audit procedures to testing complete populations. Although false positives may not entirely negate the potential benefits of testing entire populations, this paper's findings aim to provide auditors with an additional facet to consider before implementing Big Data analytics. Firms can either attempt to reduce the presence of false positives by creating sufficiently precise tests and setting higher thresholds for identifying errors (ICAEW 2016; Minniti and Camehl 2018), or they can attempt to reduce the effects of false positives. To reduce

the effects of false positives on auditors' judgments, the firm could divide the audit procedure among two auditors. The first auditor could be assigned to run the procedure, investigate each exception, identify the true errors, and remove false positives from the results. The second auditor could then evaluate the error findings. If the second auditor is blind to the fact that the procedure uncovered false positives, this would eliminate the influence of false positives on his or her judgments. Auditors can process exceptions prior to investigating them through exception prioritization (Issa and Kogan 2014); however, auditors must investigate exceptions to identify potential misstatements (Minniti and Camehl 2018). False positives can exist no matter the size of the sample, so the results of the current paper are relevant whenever sample sizes are expanded, even if the entire population is not tested.

This study is subject to limitations that provide opportunities for future research. This study provided participants with limited information. In practice, auditors will be provided with additional information about the source of the errors, why false positives were dismissed as non-errors, and other information that may influence their judgments and decisions. Additionally, in practice, auditors have to balance deadline pressures, pressures from the client, and firm pressures when making their decisions. The experimental setting is not conducive to replicating these pressures, however these pressures would likely reduce the likelihood that an auditor would recommend an adjustment. The current study is also limited by only testing certain scenarios. In the experiment, all conditions lead auditors to uncover errors. In practice, many procedures do not uncover errors, and the results of the current study may not generalize to these settings. Future research could test the effects of false positives when procedures do not

uncover any true errors. Additionally, the participants only have one source of audit evidence, thus the current paper cannot speak to how auditors weight and combine audit evidence from multiple sources, such as reperformance and inquiry. Finally, technology will play a large role in using Big Data, going forward. In the current study, I remained silent on the role of technology to equate the Sample and Population conditions. Future research can explore these limitations in order to expand the findings of the current study.

REFERENCES

Andiola, Lindsay M, Alisa G Brink, Edward J Lynch, and Jodie L Ferguson. 2019. "Client Data Files and Auditor Skepticism: How Do 'Dirty' Files Influence Auditors' Skeptical Judgments and Actions?" *Working Paper*.

Beasley, Mark S, Frank A Buckless, Steven M Glover, and Douglas F Prawitt. 2019. "EyeMax Corporation: Evaluation of Audit Differences." *Pearsons Education,* 399–404.

Beaulieu, Philip R. 2001. "The Effects of Judgments of New Clients' Integrity upon Risk Judgments, Audit Evidence, and Fees." *Auditing A Journal of Practice & Theory* 20 (2): 85–99.

Birnbaum, Michael H., and Steven E. Stegner. 1979. "Source Credibility in Social Judgment: Bias, Expertise, and the Judge's Point of View." *Journal of Personality and Social Psychology* 37 (1): 48–74. https://doi.org/10.1037/0022-3514.37.1.48.

Birnbaum, Michael H, Rebecca Wong, and Leighton K Wong. 1976. "Combining Information from Sources That Vary in Credibility." *Memory & Cognition* 4 (3): 330–36.

Braun, Karen Wilken. 2001. "The Disposition of Audit-Detected Misstatements : An Examination of Risk and Reward Factors and Aggregation Effects." *Contemporary Accounting Research* 18 (1): 71–99.

Breznitz, Shlomo. 1984. *CRY WOLF : The Psychology of False Alarms*. Englewood Hills.

Burgstahler, David, Steven M Glover, and James Jiambalvo. 2000. "Error Projection and Uncertainty in the Evaluation of Aggregate Error." *Auditing: A Journal of Practice & Theory* 19 (1): 79–99.

Burgstahler, David, and James Jiambalvo. 1986. "Sample Error Characteristics and Projection of Error to Audit Populations." *The Accounting Review* 61 (2): 233–48.

Cao, Min, Roman Chychyla, and Trevor Stewart. 2015. "Big Data Analytics in Financial Statement Audits." *Accounting Horizons* 29 (2): 423–29. https://doi.org/10.2308/acch-51068.

Cleary, Richard, and Jay C Thibodeau. 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.

Dunn, O. J. 1964. "Multiple Comparisons Using Rank Sums." *Technometrics* 6: 241–52.

Eilifsen, Aasmund, and William F. Messier. 2015. "Materiality Guidance of the Major Public Accounting Firms." *Auditing: A Journal of Practice & Theory* 34 (2): 3–26. https://doi.org/10.2308/ajpt-50882.

Elder, Randal J, Abraham D Akresh, Steven M Glover, Julia L Higgs, and Jonathan Liljegren. 2013. "Audit Sampling Research: A Synthesis and Implications for Future Research." *Auditing: A Journal of Practice and Theory* 32 (Supplement 1): 99–129. https://doi.org/10.2308/ajpt-50394.

EY Reporting. 2015. "How Big Data and Analytics Are Transforming the Audit." 2015. https://www.ey.com/en_gl/assurance/how-big-data-and-analytics-are-transforming-the-audit.

Fanning, Kirsten, Christopher P Agoglia, and M David Piercey. 2015. "Unintended Consequences of Lowering Disclosure Thresholds." *The Accounting Review* 90 (1): 301–20. https://doi.org/10.2308/accr-50861.

Glover, Steven M. 1997. "The Influence of Time Pressure and Accountability on Auditors' Processing of Nondiagnostic Information." *Journal of Accounting Research* 35 (2): 213. https://doi.org/10.2307/2491361.

Gonzalez, George C., and Vicky B. Hoffman. 2018. "Continuous Auditing's Effectiveness as a Fraud Deterrent." *Auditing: A Journal of Practice & Theory* 37 (2): 225–47. https://doi.org/10.2308/ajpt-51828.

Hayes, Andrew F. 2017. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Guilford Publications.

Hirst, D Eric. 1994. "Auditors ' Sensitivity to Source Reliability." *Journal of Accounting Research* 32 (1): 113–26.

Hurtt, R. Kathy. 2010. "Development of a Scale to Measure Professional Skepticism." *Auditing: A Journal of Practice & Theory* 29 (1): 149–71. https://doi.org/10.2308/aud.2010.29.1.149.

Hurtt, R Kathy, Helen Brown-Liburd, Christine E Earley, and Ganesh Krishnamoorthy. 2013. "Research on Auditor Professional Skepticism: Literature Synthesis and Opportunities for Future Research." *Auditing: A Journal of Practice and Theory* 32 (June): 45–97. https://doi.org/10.2308/ajpt-50361.

ICAEW. 2016. "Data Analytics for External Auditors." International Auditing Perspective. 2016. https://www.icaew.com/-/media/corporate/files/technical/iaa/tecpln14726-iaae-data-analytics---web-version.ashx.

Issa, Hussein, and Alexander Kogan. 2014. "A Predictive Ordered Logistic Regression Model as a Tool for Quality Review of Control Risk Assessments." *Journal of Information Systems* 28 (2): 209–29. https://doi.org/10.2308/isys-50808.

Kachelmeier, Steven J., and Ben W. Van Landuyt. 2017. "Prompting the Benefit of the Doubt: The Joint Effect of Auditor-Client Social Bonds and Measurement Uncertainty on Audit Adjustments." *Journal of Accounting Research* 55 (4): 963–94. https://doi.org/10.1111/1475-679X.12171.

Kahneman, Daniel, and Amos Tversky. 1972. "Subjective Probability : A Judgment of Representativeness." *Cognitive Psychology* 3 (3): 430–54.

Krahel, John Peter, and William R Titera. 2015. “Consequences of Big Data and Formalization on Accounting and Auditing Standards.” *Accounting Horizons* 29 (2): 409–22. https://doi.org/10.2308/acch-51065.

Libby, Robert, and Timothy Brown. 2013. “Financial Statement Disaggregation Decisions and Auditors’ Tolerance for Misstatement.” *The Accounting Review* 88 (2): 641–65. https://doi.org/10.2308/accr-50332.

Liberman, Varda, and Lee Ross. 2006. “Idiosyncratic Matching and Choice: When Less Is More.” *Organizational Behavior and Human Decision Processes* 101 (2): 168–83. https://doi.org/10.1016/j.obhdp.2006.01.006.

Magee, Robert P, and Mei-Chiun Tseng. 1990. “Audit Pricing and Independence.” *The Accounting Review* 65 (2): 315–36. http://www.jstor.org/stable/247627%5Cnhttp://about.jstor.org/terms.

Maines, Laureen A., and James M. Wahlen. 2006. “The Nature of Accounting Information Reliability: Inferences from Archival and Experimental Research.” *Accounting Horizons* 20 (4): 399–425. https://doi.org/10.2308/acch.2006.20.4.399.

Messier, William F., Steven M. Glover, and Douglas F. Prawitt. 2019. *Auditing & Assurance Services: A Systematic Approach*. 11th ed. McGraw-Hill.

Mileti, Dennis S., and Lori Peek. 2000. “The Social Psychology of Public Response to Warnings of a Nuclear Power Plant Accident.” *Journal of Hazardous Materials* 75 (2–3): 181–94. https://doi.org/10.1016/S0304-3894(00)00179-5.

Minniti, Robert K, and Kelen F. Camehl. 2018. *Integrating Audit Data Analytics into the Audit Process*. Association of International Certified Professional Accountants, Inc.

Nisbett, Richard E., Henry Zukier, and Ronald E. Lemley. 1981. “The Dilution Effect: Nondiagnostic Information Weakens the Implications of Diagnostic Information.” *Cognitive Psychology* 13 (2): 248–77. https://doi.org/10.1016/0010-0285(81)90010-4.

O’Hara, John M., and William S. Brown. 1991. “Nuclear Power Plant Alarm Systems: Problems and Issues.” *Proceedings of the Human Factors Society 35th Annual Meeting*, 1233–37.

Public Company Accounting Oversight Board (PCAOB). 2010a. “Audit Evidence. Auditing Standard No. 1105.” Washington, D.C.: PCAOB.

———. 2010b. “The Auditor’s Responses to the Risks of Material Misstatement. Auditing Standard No. 2301.” Washington, D.C.: PCAOB.

———. 2017. “Audit Sampling. Auditing Standard No. 2315.” Washington, D.C.: PCAOB.

Quadackers, Luc, Tom Groot, and Arnold Wright. 2014. “Auditors’ Professional Skepticism: Neutrality versus Presumptive Doubt.” *Contemporary Accounting Research* 31 (3): 639–57. https://doi.org/10.1111/1911-3846.12052.

Rose, Anna M., and Jacob M. Rose. 2003. “The Effects of Fraud Risk Assessments and a Risk Analysis Decision Aid on Auditors’ Evaluation of Evidence and Judgment.” *Accounting Forum* 27 (3): 312–38. https://doi.org/10.1111/1467-6303.00108.

Ross, Lee, Mark R Lepper, and Michael Hubbard. 1975. “Perseverance in Self-Perception and Social Perception.” *Journal of Personality and Social Psychology* 32 (5): 880–92. https://doi.org/10.1037/0022-3514.32.5.880.

Rotter, J.B. 1967. “A New Scale for the Measurement of Interpersonal Trust.” *Journal of Personality* 35 (4): 651–65.

Tetlock, Philip E., and Richard Boettger. 1989. “Accountability: A Social Magnifier of the Dilution Effect.” *Journal of Personality and Social Psychology* 57 (3): 388–98. https://doi.org/10.1037/0022-3514.57.3.388.

Tversky, Amos, and Daniel Kahneman. 1971. “Belief in the Law of Small Numbers.” *Psychological Bulletin* 76 (2): 105–10.

Vasarhelyi, Miklos A., and Fern B. Halper. 1991. “The Continuous Audit of Online Systems.” *Auditing: A Journal of Practice and Theory*.

Vasarhelyi, Miklos A, Alexander Kogan, and Brad M Tuttle. 2015. “Big Data in Accounting : An Overview.” *Accounting Horizons* 29 (2): 381–96. https://doi.org/10.2308/acch-51071.

Wickens, Christopher D., Stephen Rice, David Keller, Shaun Hutchins, Jamie Hughes, and Krisstal Clayton. 2009. “False Alerts in Air Traffic Control Conflict Alerting System : Is There a ‘ Cry Wolf ’ Effect ?” *Human Factors* 51 (4): 446–62. https://doi.org/10.1177/0018720809344720.

Wright, Arnold, and Sally Wright. 1997. “An Examination of Factors Affecting the Decision to Waive Audit Adjustments.” *Journal of Accounting, Auditing & Finance* 12 (1): 15–36. https://doi.org/10.1177/0148558X9701200102.

Yoon, Kyunghee, Lucas Hoogduin, and Li Zhang. 2015. “Big Data as Complementary Audit Evidence.” *Accounting Horizons* 29 (2): 431–38. https://doi.org/10.2308/acch-51076.

Zukier, Henry. 1982. “The Dilution Effect: The Role of the Correlation and the Dispersion of Predictor Variables in the Use of Nondiagnostic Information.” *Journal of Personality and Social Psychology* 43 (6): 1163–74. https://doi.org/10.1037/0022-3514.43.6.1163.

APPENDIX A

EXCERPTS FROM INSTRUMENT

Initial Information Screen

Population/False Positives Condition Initial Information Screen:

The major audit work in the Revenue area was a three-way match between the customer order, shipping document, and billing document. During the current fiscal year, Seven Seas sold 300 yachts, resulting in a population of 300 sales with a client provided (un-audited) book value of $452,497,200. **The complete population of 300 sales, representing 100% of the population**, was selected for testing. There were **exceptions found in 60 sales** which failed the three-way match.

The table below provides a summary of the 60 exceptions found. If all of the exceptions are true errors, the discovered misstatement would be **$2,642,555**. Remember the tolerable misstatement is $1,325,000 and overall materiality is $2,650,000.

	Count	Client Balance	Audited Balance	Difference
Exceptions	60	$90,499,440	$87,856,885	$2,462,555

Population/No False Positives Condition Initial Information Screen:

The major audit work in the Revenue area was a three-way match between the customer order, shipping document, and billing document. During the current fiscal year, Seven Seas sold 300 yachts, resulting in a population of 300 sales with a client provided (un-audited) book value of $452,497,200. **The complete population of 300 sales, representing 100% of the population**, was selected for testing. There were **exceptions found in 15 sales** which failed the three-way match.

The table below provides a summary of the 15 exceptions found. If all of the exceptions are true errors, the discovered misstatement would be **$1,331,225**. Remember the tolerable misstatement is $1,325,000 and overall materiality is $2,650,000.

	Count	Client Balance	Audited Balance	Difference
Exceptions	15	$22,624,860	$21,293,635	$1,331,225

Sample/False Positives Condition Initial Information Screen:

The major audit work in the Revenue area was a three-way match between the customer order, shipping document, and billing document. During the current fiscal year, Seven Seas sold 300 yachts, resulting in a population of 300 sales with a client provided (un-audited) book value of $452,497,200. **A random sample of 60 sales, representing 20% of the population**, was selected for testing. There were **exceptions found in 12 sales** which failed the three-way match.

The table below provides a summary of the 12 exceptions found. If all of the exceptions are true errors, the projected misstatement would be **$2,642,555**. Remember the tolerable misstatement is $1,325,000 and overall materiality is $2,650,000.

	Count	Client Balance	Audited Balance	Difference
Exceptions	12	$18,099,888	$17,571,377	$528,511

Sample/No False Positives Condition Initial Information Screen:

The major audit work in the Revenue area was a three-way match between the customer order, shipping document, and billing document. During the current fiscal year, Seven Seas sold 300 yachts, resulting in a population of 300 sales with a client provided (un-audited) book value of $452,497,200. **A random sample of 60 sales, representing 20% of the population**, was selected for testing. There were **exceptions found in 3 sales** which failed the three-way match.

The table below provides a summary of the 3 exceptions found. If all of the exceptions are true errors, the projected misstatement would be **$1,331,225**. Remember the tolerable misstatement is $1,325,000 and overall materiality is $2,650,000.

	Count	Client Balance	Audited Balance	Difference
Exceptions	3	$4,524,972	$4,258,727	$266,245

Example of Error Screen:

Exception:

Customer Number	Customer Name	Client Balance	Audited Balance	Difference
998	James Denver	$1,230,067	$1,119,361	$110,706

Click the arrow to learn the results of the investigation.

------page break---

Result: ERROR*

Example of Non-Error Screen:

Exception:

Customer Number	Customer Name	Client Balance	Audited Balance	Difference
335	Mark Batya	$1,393,092	$1,365,230	$27,862

Click the arrow to learn the results of the investigation.

------page break---

Results: NON-ERROR*

---

* In the instrument, ERROR has a red background and NON-ERROR has a green background.

Final Results Screen:

Population/False Positives Final Results Screen:

A summary of the results of the testing of Revenue is below. Exceptions labeled error exceptions were determined upon investigation to be true errors; the difference between the client balance and the audited balance for these items is shown below. Exceptions labeled non-error exceptions were determined upon investigation to not be errors; the originally identified difference was dismissed after investigation.

	Count	Client Balance	Audited Balance	Difference
Error Exceptions	15	$22,624,860	$21,293,635	$1,331,225
Non-Error Exceptions	45	$67,874,580	$67,874,580	-
Non-Exception Sample	240	$361,997,760	$361,997,760	-
Total Audited	300	$452,497,200	$451,165,975	$1,331,225

The discovered misstatement is **$1,331,225**. Remember the tolerable misstatement is $1,325,000 and overall materiality is $2,650,000.

Population/No False Positives Final Results Screen:

A summary of the results of the testing of Revenue is below. Exceptions labeled error exceptions were determined upon investigation to be true errors; the difference between the client balance and the audited balance for these items is shown below.

	Count	Client Balance	Audited Balance	Difference
Error Exceptions	15	$22,624,860	$21,293,635	$1,331,225
Non-Error Exceptions	0	$0	$0	-
Non-Exception Sample	285	$429,872,340	$429,872,340	-
Total Audited	300	$452,497,200	$451,165,975	$1,331,225

The discovered misstatement is **$1,331,225**. Remember the tolerable misstatement is $1,325,000 and overall materiality is $2,650,000.

Sample/False Positives Final Results Screen:

A summary of the results of the testing of Revenue is below. Exceptions labeled error exceptions were determined upon investigation to be true errors; the difference between the client balance and the audited balance for these items is shown below. Exceptions labeled non-error exceptions were determined upon investigation to not be errors; the originally identified difference was dismissed after investigation.

	Count	Client Balance	Audited Balance	Difference
Error Exceptions	3	$4,524,972	$4,258,727	$266,245
Non-Error Exceptions	9	$13,574,916	$13,574,916	-
Non-Exception Sample	48	$72,399,552	$72,399,552	-
Total Audited	60	$90,449,440	$90,233,195	$266,245

The projected misstatement is **$1,331,225**. Remember the tolerable misstatement is $1,325,000 and overall materiality is $2,650,000.

Sample/No False Positives Final Results Screen:

A summary of the results of the testing of Revenue is below. Exceptions labeled error exceptions were determined upon investigation to be true errors; the difference between the client balance and the audited balance for these items is shown below.

	Count	Client Balance	Audited Balance	Difference
Error Exceptions	3	$4,524,972	$4,258,727	$266,245
Non-Error Exceptions	0	$0	$0	-
Non-Exception Sample	57	$85,974,468	$85,974,468	-
Total Audited	60	$90,499,440	$90,233,195	$266,245

The projected misstatement is **$1,331,225**. Remember the tolerable misstatement is $1,325,000 and overall materiality is $2,650,000.