

## **Multidimensional Audit Data Selection (MADS) Project**

### ***1. Objective***

With the advance in data processing ability and data analytics techniques and tools, auditors no longer have to be limited to the traditional sampling approach, but can now process the entire population when performing substantive tests of details. Existing literature has introduced various statistical and machine learning techniques, which process the entire population and identify unusual transactions/data (for example, transactions that deviate from normal business activities). These are more likely to be subject to misstatement or deviation. Performing audit procedures on such questionable transactions/data increases the effectiveness of the audit by decreasing the risk of incorrect acceptance (such as concluding that the recorded account balance is not materially misstated when it is materially misstated) and the risk of incorrect rejection (in other words, concluding that the recorded account balance is materially misstated when it is not materially misstated). While these techniques are effective in helping auditors for detecting outliers (that is suspicious transactions/data), they often generate large numbers of outliers, and thus impractical for auditors to investigate in their entirety. This project aims to provide a more systematic approach for using data analytics in the audit sampling process.

### ***2. Planned Research Approach***

This project proposes a multidimensional data prioritization methodology to identify outliers, which can be suspicious transactions/data that are more likely to be problematic. To the end, we introduce a *Multidimensional Audit Data Selection (MADS)* analytics framework to guide auditors in the identification of audit data for a substantive test of details. In particular, we propose a methodology that prioritizes the identified outlier candidates based on multiple criteria prior to presenting the candidates to auditors. By allowing auditors to focus their efforts on the outlier candidates that are more likely to be at risk of misstatement or deviation, the MADS methodology reduces not only the risk of incorrect acceptance (like failing to detect material misstatement) but also the risk of incorrect rejection (such as incorrectly detecting a statement that is not materially misstated). The framework is comprised of seven components: (1) objective and criteria identification, (2) data collection and understanding, (3) data preparation, (4) MADS model building, (5) model evaluation, (6) model implementation and operation, and (7) post model evaluation. Using a real data set, we apply each component of the MADS framework to demonstrate how auditors can identify and prioritize outliers based on multiple criteria, and select a sample for a substantive test of details. Furthermore, through a survey of prior literature and by performing an empirical test, we provide evidence that our proposed approach leads to more effective audits compared to the current audit sampling approach.

### ***3. Anticipated Outcomes***

We envision that with the proposed MADS framework, auditors can systematically pin point outliers that are more likely to be subject to misstatement or deviation, and thus perform a more effective substantive test of details.

### ***4. Project timeline and milestones***

Our plan is to provide the background and support for the proposed methodology by May 2016, and provide a comprehensive guide for the MADS framework by October 2016. Finally, by December 2016, we aim to provide empirical evidence for the effectiveness of the proposed methodology.