Moving towards continuous audit and big data with audit analytics: Implications for research and practice

By Deniz Appelbaum, Alexandr Kogan, and
Miklos Vasarhelyi

of Rutgers, the State University of New Jersey

for the

35th World Continuous Auditing & Reporting Symposium

November 6&7, 2015

Hosted by: Rutgers Accounting Research Center (RARC) and Continuous Auditing & Reporting Laboratory (CAR-Lab)

Discussion of the current external audit environment

"Advances in technology and the massive proliferation of available information have created a new landscape for financial reporting. With investors now having access to a seemingly unlimited breadth and depth of information, the need has never been greater for the audit process to evolve by providing deeper and more relevant insights about an organization's financial condition and performance –while maintaining and continually improving audit quality.

- Does this mean that core elements of the audit such as the current "pass/fail opinion" that external auditors are mandated to provide – and that has served investors well for years, need to expand? Absolutely!" (Liddy, 2015)

(James P. Liddy is KPMG LLP U.S. Vice Chair, Audit and Regional Head of Audit, Americas. Article published in Forbes August 4, 2015.)

Discussion of the current external audit environment

Issues facing the current external audit:

- Should new (modern) analytics methods be used in the audit process?

- Which of these methods are the most promising?

- Where in the audit are these applicable?

- Should auditing standards be changed to allow / facilitate these methods?

- What are the competencies needed by auditors in this environment?

Audit Data Analytics (ADA): One way to define…..

Audit Data Analytics (ADA) is the analysis of data underlying financial statements, together with related financial or non-financial information, for the purpose of identifying potential misstatements or risks of material misstatement.
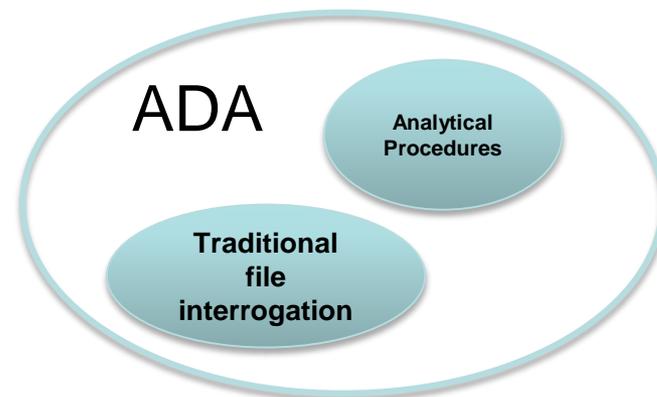
ADA includes methodologies for:

- Identifying and analyzing anomalies in the data
- Identifying and analyzing patterns in the data including outliers
- Building statistical (e.g., regression) or other models that explain the data in relation to other factors and identify significant fluctuations from the model
- Synthesizing pieces of information from disparate analyses and data sources into wholes that are greater than the sum of their parts for purposes of overall evaluation

ADA defined in this way *includes:*

- Analytical Procedures (AU-C 520)—preliminary, substantive, and FS review—including reasonableness testing
- Traditional file interrogation

ADA

Analytical Procedures

Traditional file interrogation

(Liu, 2014)

ADA *mode* can be *exploratory* or *confirmatory*

	Exploratory mode	Confirmatory mode
When	Planning	Performance
Question	What is going on here? Does the data suggest something might have gone wrong? Where do the risks appear to be? What assertions should we focus on?	Does the data conform with and thus confirm my model for what ought to be?
Approach style	Bottom-up, inductive, few starting assumptions, assertion-free	Top-down, deductive, model-driven, starts with development of model based on assertions to be tested
Methods	Graphical visualizations used to discover patterns in and understand the data—possibly several to get different viewpoints	Comparison of actual data to model taking into account materiality, desired assurance and assertions being tested; more mathematical than graphical
Results	Identified risks, areas of focus, potential models for confirmatory stage	Identified anomalies, unexpected patterns, outliers and other significant deviations

(Stewart, 2015)

ADA Examples

Exploratory

- Cluster analysis
- Text and data mining
- Data visualization
  - Scatterplots
  - Scatterplot matrices
  - Line charts
  - Spread charts
  - Needle graphs
  - Small multiples of graphics
  - Heat maps
  - Treemaps
  - Relationship maps

Confirmatory

- Analytical procedures
  - Regression analysis, ratio analysis
  - Reasonableness tests
- Recalculations
- Traditional file interrogation
  - Footing, extending
  - Duplicate detection
  - Out-of-range detection
  - Other 100% tests
- Journal entry testing (SAS 99)

Exploratory and confirmatory ADA is a spectrum of analytics and the processes are iterative, starting with exploratory

4

(Stewart and Liu, 2014)

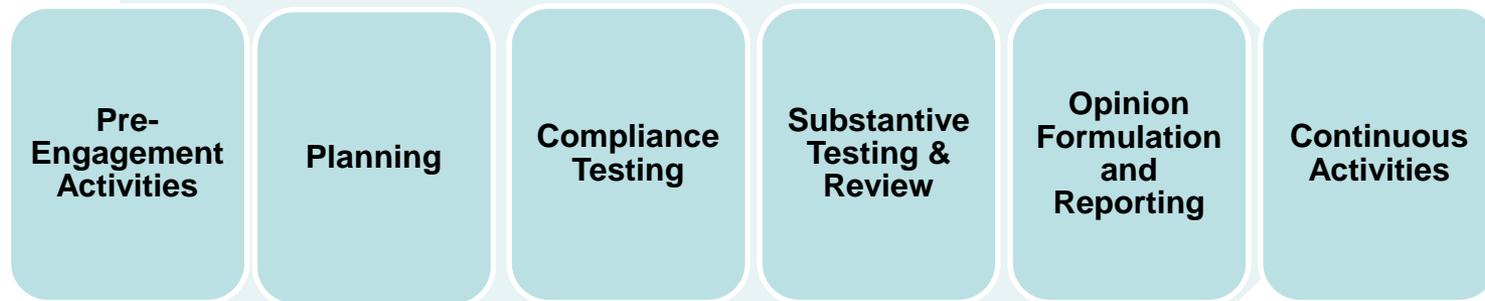# Should auditing standards be changed to facilitate these methods?

- The economics of the audit is constrained by PCAOB requirements
- Sampling requires laborious follow ups but in a population of millions or hundreds of thousands there is little meaning on picking 25 transactions and reviewing them.
- For example three way matches used to be performed manually and reviewed manually. Now advanced accounting systems and ERPs perform these automatically. Is this performance audit evidence, new analytics, or just automation? If automation how do the audit standards take this into consideration? Is there a difference between automation and analytic methods?
- The new forms of evidence may use external and internal data (Brown-Liburd & Vasarhelyi, 2015) potentially from external sources like social media but providing valuable tertiary audit evidence that may be used to complement / replace current tests. Would these need new guidance?

What are the competencies needed by auditors in this environment?

- Accounting faculty tend not to be prepared to teach analytics
- There is a widespread feeling that students are not interested
- The accounting curriculum is very full to add more IT, statistics, and modeling
- As the CPA exam does not include these topics, there is little motivation by students for their addition to the curriculum of study
- Firms will tend / or already have hired specialist groups from non-accounting backgrounds
- Practitioners are also not prepared and their internal audit practices have not caught up properly with these issues.
- Firms have been developing software to improve their processes but feel curtailed by the PCAOB examination process

RUTGERS

External Audit Analytics Framework

- Framework is defined by three Business Analytics dimensions (Holsapple et al, 2014): Domain, Orientation, Technique
- Domain: Audit model of Loebbecke and Cushing (Elliott, 1983) is the context for the analytical procedures

| Pre-Engagement Activities | Planning | Compliance Testing | Substantive Testing & Review | Opinion Formulation and Reporting | Continuous Activities |

External Audit Analytics Framework

Orientation: descriptive, predictive or prescriptive

Techniques:

- qualitative or quantitative
- deterministic or statistical
- based on unstructured, semi-structured, or structured data
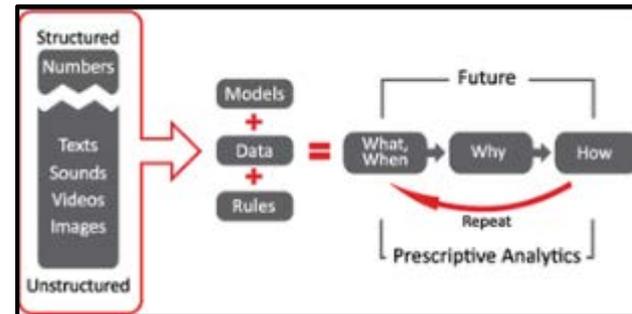
The most commonly used AA techniques are those that are quantitative, statistical, and based on structured data

The dominance of quantitative techniques in AA is due to the fact that the main objective of external audit is to provide assurance on the accounting numbers. Therefore, the accounting numbers are the quantities that are the focus of AA.

*(Wikipedia, 2015)*

External Audit Analytics Framework

- **Expectation Models:** an empirical relationship among several accounting numbers or some other important quantitative measures of business operations and is inferred from the archive of historical records

- **Structural techniques** look for various structural properties in the historical records. A very popular recent example is process mining

- The primary objective of **multivariate techniques** is to develop relationships between or among variables/features under study

- **Latent features:** common classification and regression techniques do not work explicitly with any latent features, while common **clustering** techniques do. For example, while it is commonly assumed that fraud is a routine occurrence in most enterprises (though, most likely, at a low level), very few confirmed and documented cases of fraudulent transactions exist. **The challenge? - creating expectation models for what is fraudulent versus normal based on the historical record that does not identify past transactions in this way**

Literature Review of Analytical Procedures in the External Audit

- Evidence collection and analytical procedures - Both processes are required by the standards, the degree of which is according to auditor professional judgment

- The standards do not explicitly define the type of analytical approaches that should be undertaken by public auditors to fulfill these regulatory requirements, except that the auditor should develop an expectation from the appropriate analytics of reliable data from certain accounts, and then calculate the difference of these expectations and the recorded numbers

- The standards require that analytical procedures be undertaken in addition to evidence collection at the preliminary review and final review stages (Daroca & Holder, 1985), but the analytical approaches are left to auditor judgment

- To date, 222 papers have been classified as discussing analytical procedures in the context of the external audit model of Loebbecke/Cushing

Literature Review of Analytical Procedures in the External Audit

- Initial keyword search of Audit papers database (Trotman et al, 2009): *"analytics", "analytical review", "audit planning", "risk assessment", "internal control assessment", "compliance testing", "substantive testing activities", "review", "fraud", "Going Concern",* and *"Fair Value Assessment"*

- Papers were categorized according to audit model phase, substantive risk procedures, IC assessment procedures, research approach/method, AA techniques, auditor qualities, assertions, keywords, journal, abstract and results
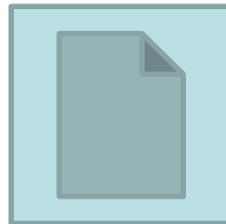
Literature Review of Analytical Procedures in the External Audit

Number of Papers per Audit Phase
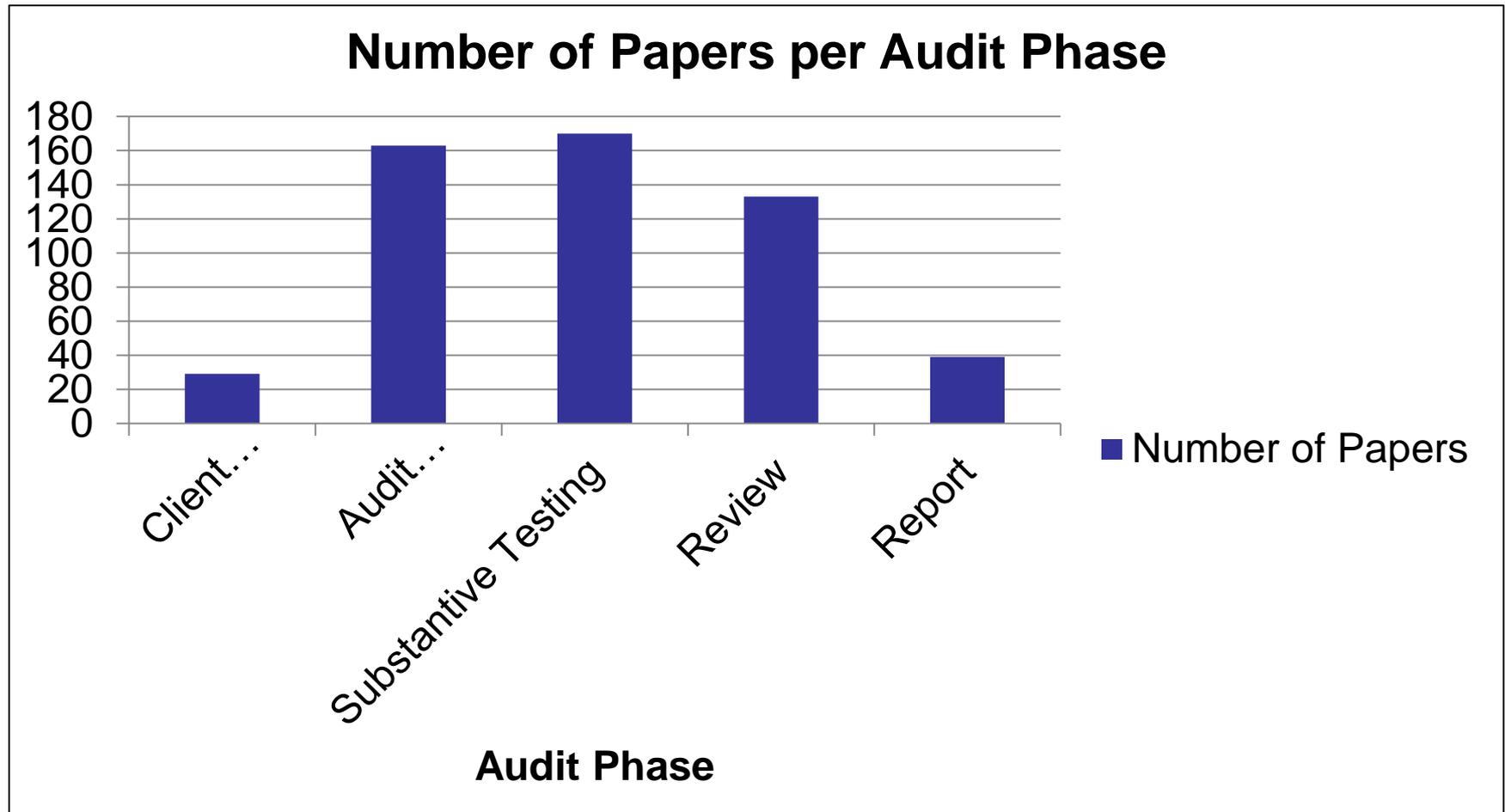
Literature Review of Analytical Procedures in the External Audit : Techniques

Machine Learning/Data Mining:	Process Mining	Simulation, Process Optimization
	Advanced Classifiers	Support Vector Machine (SVM), Artificial Neural Networks (ANN), Multilayer Feed Forward Neural Network (MLFF), Genetic Algorithm
	Rules-Based Classifiers	Expert Systems/Decision Aids, Majority Vote, AntMiner +
	Ensemble Methods	Boosting, Bagging, Bootstrap
	Decision Trees	C4.5 statistical classifier
	Bayes Classifiers/Probability Models	Bayesian Theory/Bayesian Belief Networks (BBN), Naïve Bayes, Dempster-Shafer Theory, Probability Theory
	Other	Clustering, Text Mining, Visualization, Group Method of Data Handling (GMDH)

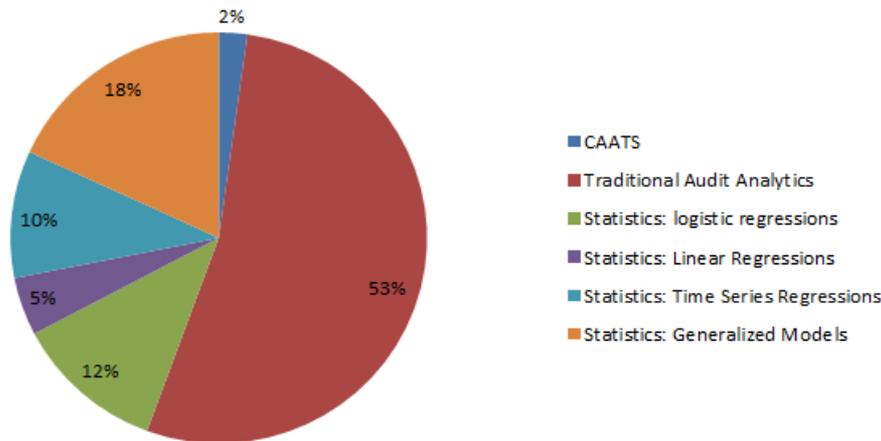# Literature Review of Analytical Procedures in the External Audit : Techniques

"Traditional" Audit Analytics:	CAATS	Transaction Tests, Data Modeling, Data Analytics
	"Traditional" Analytical Review/Analytics	Sampling, Ratio Analysis, Firm developed proprietary software
	Statistics: Log Regression	Log Regression, Step-Wise Logistic, Ordinal Regression Model
	Statistics: Linear Regression	Linear Regression
	Statistics: Time Series Regressions	Time Series Regression, Auto Regressive Integrated Moving Average (ARIMA), Box Jenkins (ARIMA), Random Walk (ARIMA), Random Walk Drift (ARIMA), Seasonal Time Series X-11, Martingale, Sub-Martingale, Single and Double Exponential Smoothing Model
	Statistics: Generalized Models	Multicriteria Decision Aid, Multivariate Distribution, Benford's Law, Descriptive Statistics, Univariate and Multivariate Regression Analysis, Structural Model, Analytical Hierarchy Process (AHP), Spearman Rank Correlation Measurements, Complementary Hypothesis Evaluation, Independent Hypothesis Evaluation, Monte Carlo Study/Simulation

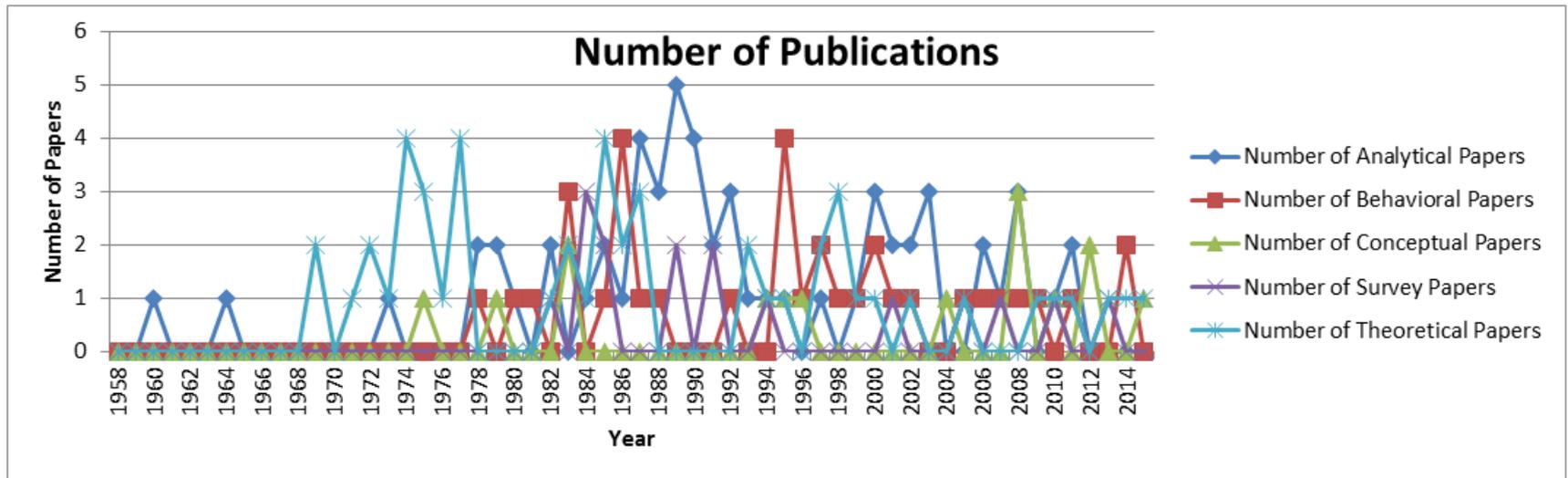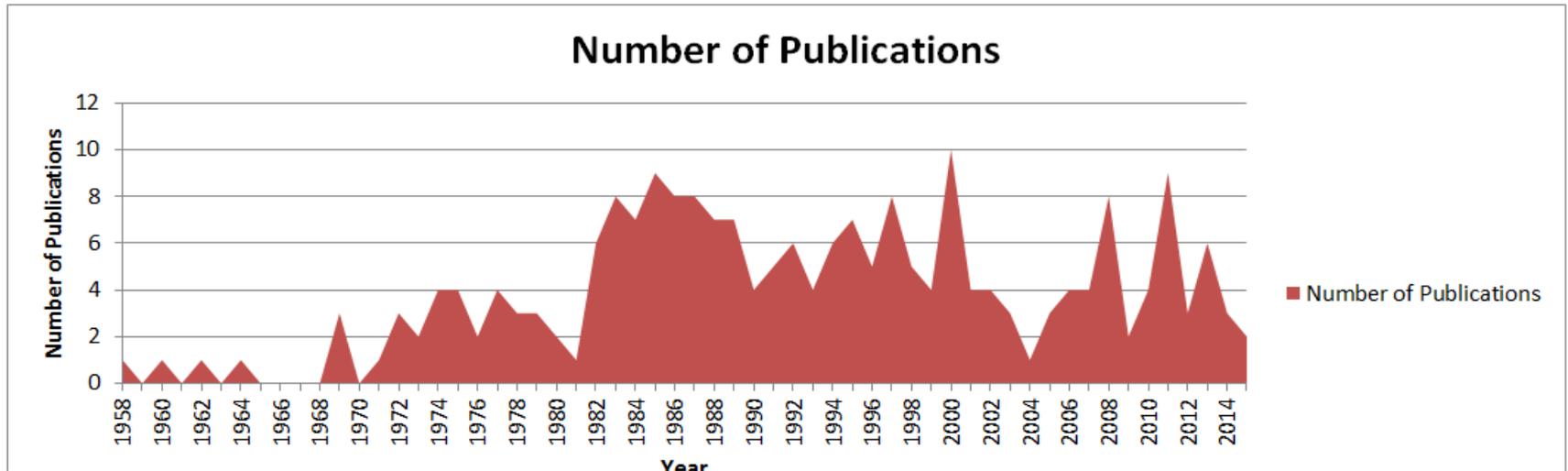# Literature Review of Analytical Procedures in the External Audit

Percentage of papers using Machine Learning/Data Mining

- Process Mining — 11%
- Advanced Classifiers — 11%
- Rules Based Classifiers — 20%
- Ensemble Methods — 1%
- Decision Trees — 2%
- Bayes Classifiers/Probability Models — 45%
- (additional) — 10%
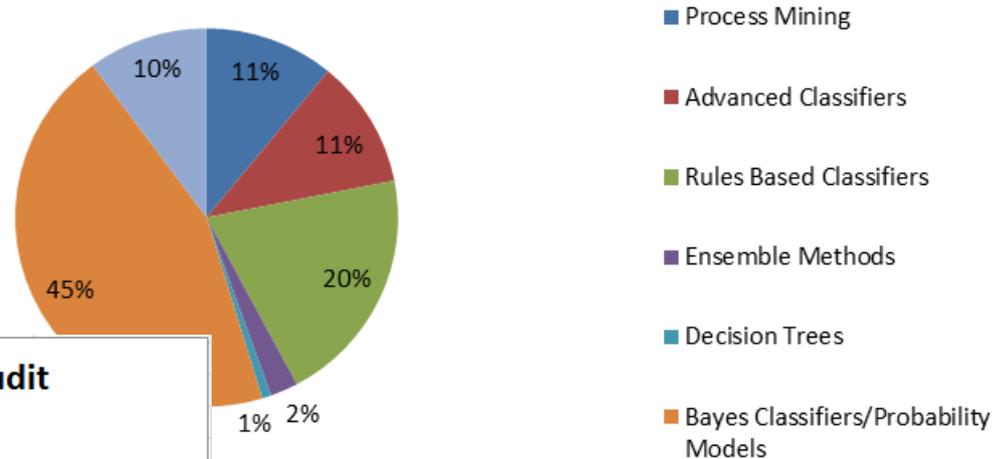
Percentage of Papers using Traditional Audit Analytics

- CAATS — 2%
- Traditional Audit Analytics — 53%
- Statistics: logistic regressions — 12%
- Statistics: Linear Regressions — 5%
- Statistics: Time Series Regressions — 10%
- Statistics: Generalized Models — 18%

17

Literature Review of Analytical Procedures in the External Audit

Number of Papers per Journal

Legend: ■ Number of Papers
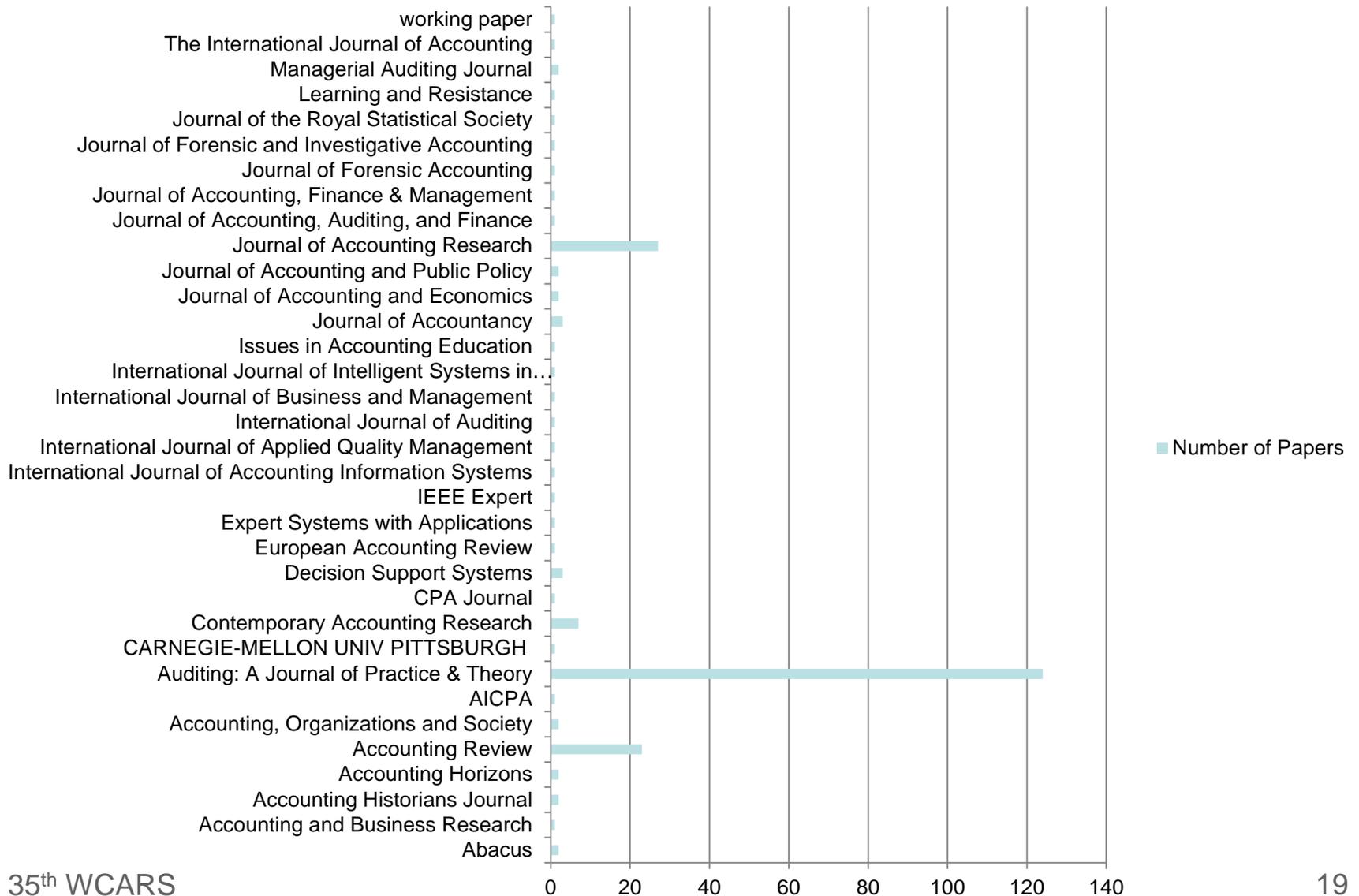
# Literature Review of Analytical Procedures in the External Audit

- Sampling: "the application of an audit procedure to less than 100 percent of the items within an account balance or class of transactions for the purpose of evaluating some characteristic of the balance or class." (PCAOB AU 350.01)

- AU-C 330.A65-71 (AICPA 2011): testing can be (1) 100 percent/all items, (2) selecting specific items, and (3) audit sampling.

Number of Papers per Year Discussing Sampling
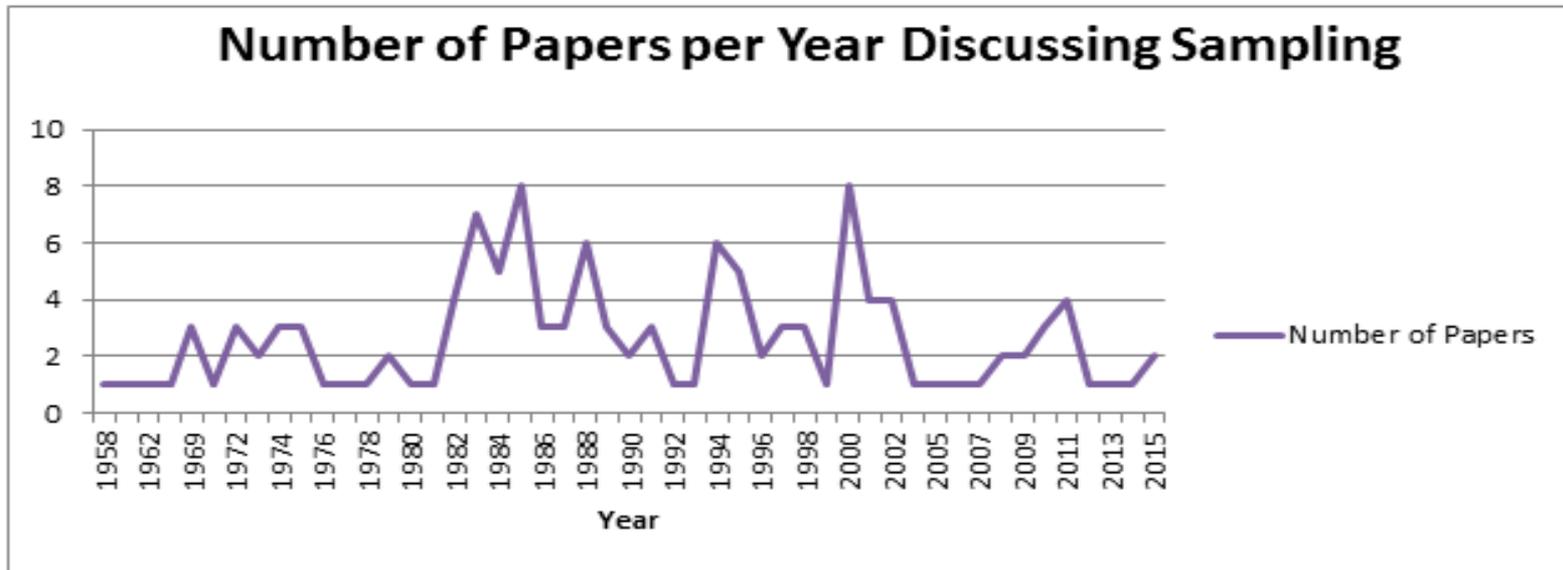
Prescriptive Audit Analytics: Looking Forward

- "It has also been shown that many internal audit procedures can be automated, thus saving costs, allowing for more frequent audits and freeing up the audit staff for tasks that require human judgment (Vasarhelyi, 1983, Vasarhelyi, 1985; Alles, Kogan, and Vasarhelyi, 2002)." (AICPA, 2015)

- Audit Methods have been retroactive as most manual methods relied on some degree of manual verification of source documents or third party verification of balances thru manual confirmation

- Overall, the expected value of assurance efforts must be larger than its costs. Once manual efforts are voluminous they become very expensive.

Evolving the Environment
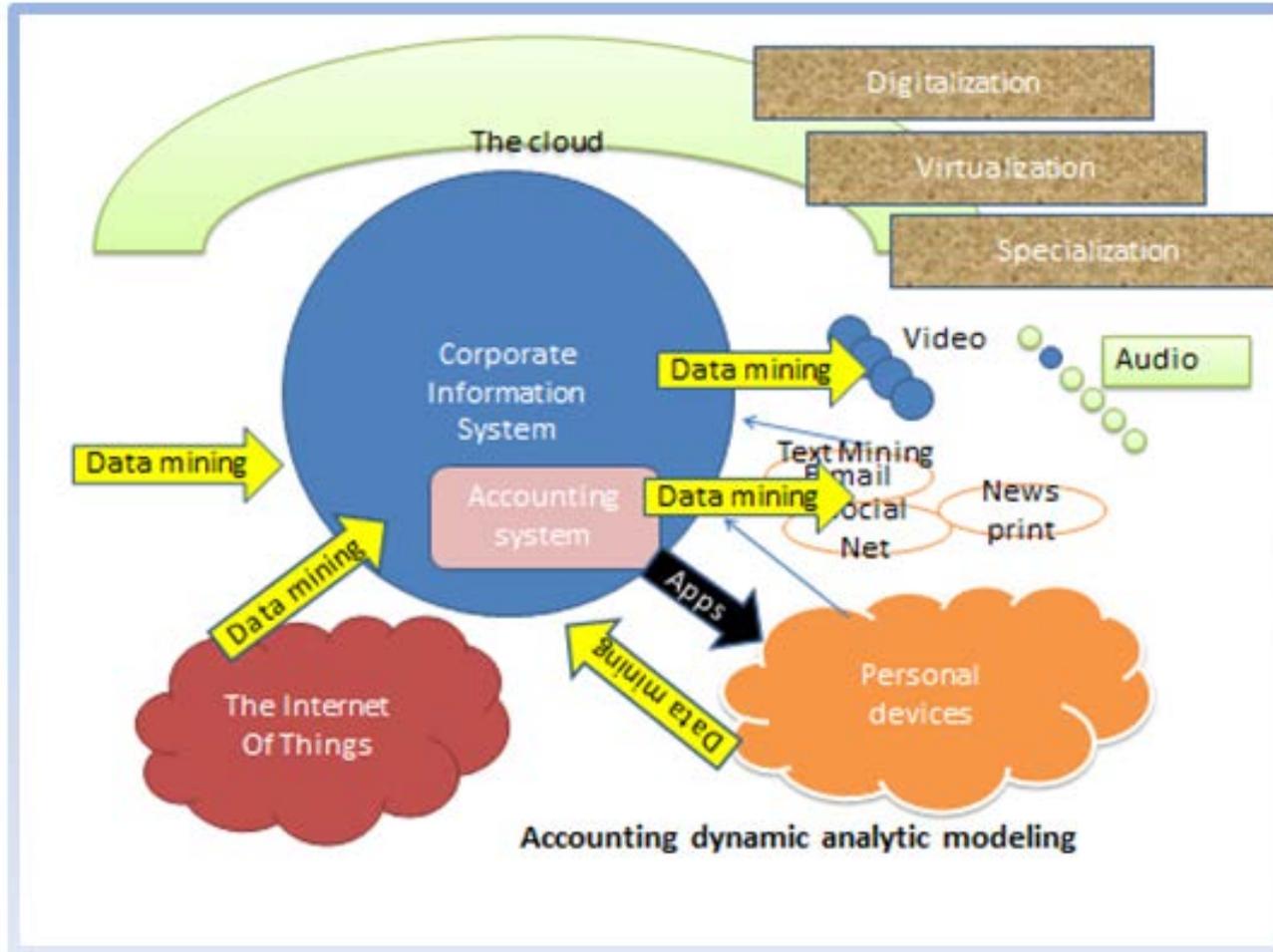
The evolutionary environment (adapted Liu and Vasarhelyi, 2014)

An Audit Eco-System
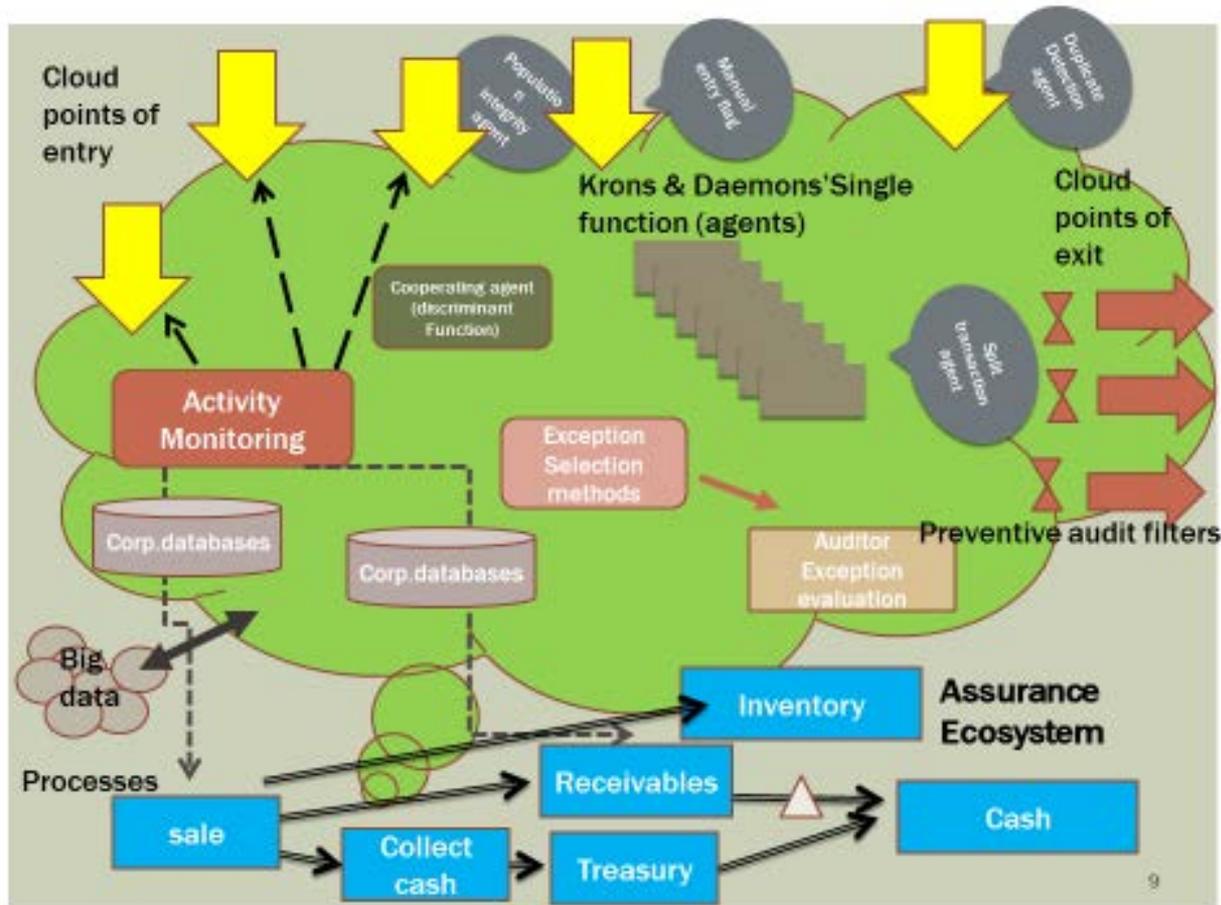
Audit Ecosystem (adapted from AICPA, 2015 chapter 1)

Some research questions

- The appropriateness of the method for a particular audit function?

- How should the assurance function be reorganized to better use ADA?

- How can predictive technologies be used to set comparison models against which match actuals. How to set allowable variance (Vasarhelyi & Bumgartner, 2015)?

- How to set the timing of performing an assurance function, agents may run very frequently or intermittently?

- What additional verification processes would be desirable with the extant analytic technology?

- What parts of the audit standards and processes must be progressively abandoned due to their obsolescence?

- What types of "suspicion functions" should be set up for a preventive audit or just for transaction or account review?

- How can validation function be developed that link corporate information with big data variables to validate the dimensionality and predict variances?

- How can we migrate to a complex ecosystem and more advance assurance processes while not disrupting the current processes?

Discussion/Conclusion

- Big data and business analytics have dramatically changed the business environment and the capabilities of business processes.

- Business functions are changing, business capabilities are being added, anachronistic business functions are being eliminated, and most of all, processes are being substantially accelerated.

- The same has to happen to the assurance function, its rules need to be changed, its steps evolved, automation integrated to its basic processes, its audit analytical procedures enhanced, and its timing close to real-time in predictive and preventive modes