A Vetting Protocol for the Analytical Procedures Platform for the AP-Phase of PCAOB Audits

Mohamed Gaber[1] & Edward J. Lusk[1]

[1]The State University of New York (SUNY) at Plattsburgh, Plattsburgh, NY, USA

Correspondence: Edward J. Lusk, the State University of New York (SUNY) at Plattsburgh, Plattsburgh, NY, USA.

Abstract

Study Context AS5[2017], issued by the Public Company Accounting Oversight Board, requires the use of Analytical Procedures [AP] at the Planning and Substantive Phases of Assurance Audits of firms traded on active exchanges. Logically, an aspect of this requirement is satisfied by using a Panel of the Client's data at the Planning Phase to <u>forecast</u> the Client's YE-closing values and then at the Substantive Phase to dispose the directional difference between the: [Actual Client's YE-value and the AP-Forecasted YE-value]—the Disposition Phase. **Research Focus** To date, neither the PCAOB nor the AICPA have suggested a pilot-test paradigm to vet the AP-forecasting Protocol under consideration. To address this lacuna, we detail an AP: Decision Support System [AP:DSS] that offers to the Audit InCharge a two-stage pre-analysis AP-vetting [Pilot-Test] platform that employs False Negative [FN] and False Positive [FP] Profilers. In inferential analyses, the FP-Risk is usually benchmarked using the FN-Risk. **Deliverables** A comprehensive AP-vetting model is offered and illustrated using: (i) a preliminary estimator of a reasonable sample size, (ii) two Standard Forecasting Models: The *Excel* versions of the OLS Linear Two-parameter and the Moving Average Models, and (iii) a Benchmarking protocol. Unique in this AP:DSS vetting protocol is that the FP-risk is contexted by the FN-risk from the independent benchmark domain. This duality enhances the inferential impact of the vetting protocol as it uses separate variable sets. The AP:DSS is available at no cost as an e-Download.

Keywords: forecasting, pilot-testing, false negative, false positive profiling

1. Forecasting as the Principal Platform in Forming an Analytical Procedures [AP] Investigation Projection

1.1 Context

The motivation for this research report is to offer: (i) an AP-forecasting-model paradigm(Note 1) to <u>forecast</u> at the Planning Phase the Client's YE-closing values and then at the Substantive Phase to dispose the directional difference between the: [Actual Client's YE-value and the AP-Forecasted YE-value] called: the AP[P-S]:Protocol, and (ii) an inferential pilot-test suitable for vetting the AP[P-S]:Protocol before it is employed in the audit. Such an inferential accoutrement seems long overdue as, to date, neither the PCAOB nor the AICPA have suggested a protocol to form inferential vetting information re: the AP[P-S]: Protocol. This is surprising as the AICPA (2012) has provided an excellent illustrative case called the *On the Go Stores* where a forecasting model was detailed; however, there was no pre-test vetting suggested.

To investigate and detail enhanced forecasting modeling choices and to create the inferential basis of their evaluation will require the consideration of the specific logistical design aspects of the AP-[P-S]: Protocol. The central feature of these considerable design challenges is reliance on the professional judgment of the auditor. Following, these aspects will be considered.

1.2 The Ensemble of the Planning and the Substantive Phases of the Audit

The linkage of these two AP-phases is concisely summarized by Arens, Elder, Beasley & Hogan (2015, p.193) as:

> "Analytical procedures are *required* at the *planning phase* as part of risk assessment procedures to understand the client's business and industry and to assist in determining the nature, extent, and timing of audit procedures. - - - . Analytical procedures are also *required* during the *completion phase* of the audit. Such tests serve as a final review for material misstatements or financial problems and help the auditor take a final "objective

> look" at the audited financial statements. - - -. Analytical procedures performed during audit planning generally use aggregated data to help understand the client's business and identify areas where misstatements are more likely. In contrast, substantive analytical procedures used to provide audit evidence require more reliable evidence, often using disaggregated data for the auditor to develop an expectation of the account balance being tested."

This linkage between the AP-Planning and the AP-Substantive phases all but proscribes that there would be a forecasting projection at the Planning stage. Typically, these forecasts are formed analytically using a forecasting model with inputs from the previously reported and certified past data drawn from the client's Accounting Information System [AIS]. At the Substantive Phase these forecasts are compared to the client's Year-End value. As offered by O'Donnell & Perkins (2011), this comparison is called the AP-Disposition Phase and produces critical information that suggests the nature of the Risk of the audit. This is used by the InCharge to re-calibrate the Audit Risk and so enters into consideration of the need and nature of Extended Procedures Investigations [EPI] at the Substantive Phase. These procedural linkages are consistent with the work of Wheeler and Pany (1990) and Fukukawa, Mock & Wright (2011) who also emphasize that AP are critically important in calibrating the risk-level of audit so that the auditor can use this risk-assessment to intelligently design the testing procedures for collecting audit evidence in the best practices execution of the audit. Finally, to better understand the historical underpinnings of the importance of the AP-investigations in the audit context, a must read, is the advice offered by Bettauer (1975); this AP-context as a driver for the audit was identified many decades ago.

The modelling form used in the AP[P-S]:Protocol is very often the standard linear regression model. In fact, as noted above, the AICPA (2012) initiative: The Clarity Project dealing with Analytical Procedures presents a carefully detailed study called the *On the Go Stores* case where the forecasting technique used is the standard linear regression model.

1.3 Judgment: The Decision-Making Imperative

Let us consider first the judgment aspect of the AP[P-S]:Protocol. The lynch-pin that rationalizes the quality of the audit is the judgement of the InCharge. There are <u>two</u> authoritative references that focus on the critical nature of judgment in the conduct of the audit. The <u>first</u> is given by the AICPA: Generally Accepted Audit Standards [GAAS (Note 2)]:

> *GAAS:GS1 The auditor must have adequate technical training and proficiency to perform the audit.*
>
> *GAAS : SFW2 The auditor must obtain a sufficient understanding of the entity and its environment, including its internal control, to assess the risk of material misstatement of the financial statements whether due to error or fraud, and to design the nature, timing, and extent of further audit procedures.*

To coordinate with the GAAS, the <u>second</u> authoritative source mentioned above is the PCAOB provides important elaborations and guidance relative to AS5[15Dec2017] wherein it is emphasized that the Judgement of the InCharge is a critical feature underlying the PCAOB-Audit. It is most instructive, in fact these sections are required reading in our Auditing and Assurance course, to consider the following comments taken from the current PCAOB rule set:

AS 1001: Responsibilities and Functions of the Independent Auditor:

> Page :5 : 05. *In the observance of the standards of the PCAOB, the independent auditor must exercise his judgment in determining which auditing procedures are necessary in the circumstances to afford a reasonable basis for his opinion. His judgment is required to be the informed judgment of a qualified professional person.*

AS 1010: Training and Proficiency of the Independent Auditor:

> Page: 9: 03. *In the performance of the audit which leads to an opinion, - - -. The junior assistant, just entering upon an auditing career, must obtain his professional experience with the proper supervision and review of his work by a more experienced superior. The nature and extent of supervision and review must necessarily reflect wide variances in practice. The engagement partner must exercise seasoned judgment in the varying degrees of his supervision and review of the work done and judgments exercised by his*

subordinates, who in turn must meet the responsibilities attaching to the varying gradations and functions of their work.

AS 2305: Substantive Analytical Procedures:

> Page9:09. *The auditor's reliance on substantive tests to achieve an audit objective related to a particular assertion may be derived from tests of details, from analytical procedures, or from a combination of both. The decision about which procedure or procedures to use to achieve a particular audit objective is based on the auditor's judgment on the expected effectiveness and efficiency of the available procedures. For significant risks of material misstatement, it is unlikely that audit evidence obtained from substantive analytical procedures alone will be sufficient. (See paragraph .11 of AS 2301, The Auditor's Responses to the Risks of Material Misstatement.)*

AS 2315: Audit Sampling:

> Page.207:01. *Audit sampling is the application of an audit procedure to less than 100 percent of the items within an account balance or class of transactions for the purpose of evaluating some characteristic of the balance or class. This section provides guidance for planning, performing, and evaluating audit samples.*
>
> *.02 The auditor often is aware of account balances and transactions that may be more likely to contain misstatements. He considers this knowledge in planning his procedures, including audit sampling. The auditor usually will have no special knowledge about other account balances and transactions that, in his judgment, will need to be tested to fulfill his audit objectives. Audit sampling is especially useful in these cases.*
>
> *.03 There are two general approaches to audit sampling: nonstatistical and statistical. Both approaches require that the auditor use professional judgment in planning, performing, and evaluating a sample and in relating the evidential matter produced by the sample to other evidential matter when forming a conclusion about the related account balance or class of transactions. Either approach to audit sampling can provide sufficient evidential matter when applied properly. This section applies to both nonstatistical and statistical sampling.*

These GAAS and PCAOB guidance principles rationalize that the judgmental decisions of the InCharge auditor are the basis of the Assurance Audit. The important implication of this is:

As the Basis of the collection of Audit Evidence is a Random Sample of Sufficient size from the Population of Sensitive AIS accounts under Audit Scrutiny, the justification of the Opinions scripted in PCAOB Audits is NOT if the evidence is generated by an Error-free audit protocol rather it is:

If the InCharge has exercised judgement born of training and experience in developing a relevant, reliable, and inferential tested audit protocol, this, we offer, is the due diligence test of the quality of the Audit.

1.4 Study Précis

This is the point of departure for our study. Following, we wish to extend the forecasting context illustrated by the AICPA:Clarity Project by suggesting a (i) number of technical enhancements in the form of a Decision Support System [DSS], called: AP:DSS, intended to rationalize or justify the particular AP[P-S]:Protocol selected by the InCharge, and (ii) provide an inferential basis for its evaluation. Otherwise said: We will suggest an inferential Pilot Test for the particular AP[P-S]:Protocol under consideration that is based upon a DSS called: AP:DSS. The AP:DSS is parameterized by the auditor's judgment and generates an information profile that may be used by the InCharge to vet the AP[P-S]:Protocol. To achieve our goal we will do the following:

1. Detail the context for **Integrating Forecasts** into the AP[P-S]:Protocol of the audit,

2. Provide an overview of the pre-analysis **Vetting** of the forecasting protocol of the AP[P-S]:Protocol,

3. Offer a **Sample Size profiler** for initializing the AP:DSS,

4. Give the details of the **AP:DSS** used to automate the creation of information for Vetting-Pilot Testing the AP[P-S]:Protocol,

    i. Offer **Two Excel Forecasting Models** that are ideal for creating AP YE-projections,

ii.   Introduce the **False Positive and False Negative Error Measures** that are needed to make an intelligent decision on the inference of evaluating the forecasting dimension of the AP phase of the audit,

2. Forecasting: An Integral Aspect of the Analytical Procedures

2.1 Contextual Focus: Forecasting in the Service of Analytical Procedures

The variations in the AP configurations that one finds in practice indeed spans the spectrum; this is likely due to the fact that nowhere in AS2, AS5[2007:2017] are specifics given as to the technical modalities that would constitute a "best-practices" configuration regarding the execution of the AP as linked from the Planning Phase through the Substantive Phase. This is most likely the reason that the AICPA offered the *On the Go Stores* illustrative example.

2.2 Normative Inference Protocol for AP-Forecasting

As Gaber & Lusk (2018) note single projections as to the YE-value have very little value; thus, it is the usual case that projection made at the Planning Phase should have capture confidence intervals so that the client-YE value can be evaluated in the binary profile context: The YE-value is IN the interval or NOT in the interval. This percentage profile is then used by the InCharge to draw an inference that will be used in making the decision re: launching an Extended Procedures Investigation [EPI]. As *statistical inference* is the driver of using forecasting capture intervals in the AP-context and given the obvious wisdom of vetting or pilot-testing the AP[P-S]:Protocol, we propose that the basis of the vetting protocol of the AP[P-S] protocol should be *inferential*. The idea is simple: The InCharge will test the AP[P-S] forecasting configuration that is expected to be used to generate the inferences that will be used to create audit evidence to determine *a priori* if the AP[P-S]:Protocol passes the "reasonability" inferential test. If not, other AP[P-S]:Protocol configurations may be tested. This pre-testing is needed to avoid, insofar as possible, the Type III Error—to wit selecting the wrong or inappropriate inference protocol. Without such a pre-employment vetting there will be an issue as to whether the inferences were likely to have been affected by an inappropriate selection of the AP-forecasting configuration. In our experience, both for assurance audits and also in internal audits, pilot-test-vetting is rarely used. Consider following the set of information that is needed at the vetting phase.

3. Vetting Context for the Forecasting in the AP[P-S] Context

3.1 Pre-Analysis: Analytic Common-Sense

Initially, it is useful to define the concept of Vetting. Paraphrasing the National Vetting Center (Note 3) we offer that Vetting—as an inferential context is: Providing a clearly related sampling frame that is: (i) independent of the final or actual accrual, but (ii) logically relevant to the inferential problem. The benefit of a pre-analysis vetting it to improve the confidence in the inferential results. In this sense, if a protocol is vetted this is akin to a Rosenthal (1991) meta-analysis—i.e., there is more confidence in the vetted result relative to (iii) to the inferential context. Most simply: *Look Before You Leap*.

The quality of any inferential protocol is how attuned it is to the "population" reality which is assumed to underlie the empirical data. Assuming that the population is the correct reference-sampling frame, the next issue in vetting is to identify a vetting frame that "mirrors" the actual AP-sampling frame. This is important, as there is a conceptual issue with sampling from the same population to vet the testing protocol. This means that a benchmark dataset is usually needed to provide a rich vetting context.

3.2 Parametrizing the AP[P-S] Proposed Protocol

To begin the development of the required information for vetting the AP[P-S] protocol, we will offer the set of parameters that should be generated by the audit-team in a typical non-forensic continuing audit of our experience (Note 4).

At the Planning Phase meeting, after the PCAOB required *Walkthrough* [PCAOB:15Dec2017 (Note 5)] has been completed, **the InCharge will collect opinions on**:

In the context of the AP[P-S]:Protocol, how often, for the sensitive accounts [discussed subsequently] to be tested, would an InCharge likely assume that management's system of internal control over financial reporting [ICoFR] was adequate [or not] to the extent that a finding of a 302-Disclosure Weakness (Note 6) would be [or not be] warranted? To answer this generalized musing, a specific testing protocol AP[P-S]:Protocol must be developed, parametrized and inferentially queried. For our illustrative, we have selected the following from the many possible parameterizations:

What is the [maximum percentage of time], Ho in the inferential context, for the forecasting model selected for the AP[P-S]:Protocol-disposition phase (Note 7) *that if the YE-values were to be in the Confidence/Capture Intervals of the Forecasting model, that the InCharge would begin to suggest the need of an EPI of the client's AIS re: ICoFR?* Expressed in other words:

What is the forecasting CI-capture-percentage that is the decision-making binary-partition at or below which the InCharge would likely effect an EPI, but above which would not effect an EPI. This is the Ho-percentage for inferential purposes. This is a directional test for Ha > Ho where if Ho is rejected that leads to the conclusion that EPI are in fact Not warranted? Simply: If this alternative percentage, Ha, were to be the Actual State of Nature and be found to be inferentially > Ho, then the InCharge would reject, Ho, and believe that there were no issues with the ICoFR as the capture percentage is more than implied by Ho i.e., the capture percentage in the CIs is higher than Ho, and thus that there would likely be no issues with clients ICoFR, and in this case the InCharge would not consider launching an EPI investigation of the Client's AIS re: ICoFR?

Assume that the InCharge, after discussion at the Planning Stage with the audit team, fixes the following parametrizing values for the AP[P-S]:Protocol:

Ho = 45% of the client's YE-Account are IN the forecasting capture/confidence interval. Assumed Indication: The Client's ICoFR is not adequate and so EPI would be warranted. Recall, this is the largest percentage to be in the capture intervals of the forecasting model used in the AP[P-S]:Protocol where the InCharge will nonetheless launch an EPI.

Alternative for Inferential Testing: Ha = 50% of the client's YE-Account are IN the forecasting capture/confidence interval. Assumed Indication: The Client's ICoFR IS adequate and so EPI would not likely be warranted. As *Ha* is > *Ho* and as *Ho* is the binary toggle-point for the EPI-decision, then if 50% were to be the State of Nature then *Ho* would be rejected and EPI would not be effected.

Note should be taken that this is one of many possible parameterizations. In addition, the InCharge has selected the case where the YE-values are in the capture interval; alternatively, the InCharge could have selected the percentage of time that the forecasting confidence intervals did not capture the YE-value. Either is possible, as this choice-variant will just condition the inference of the vetting test. This is to say that in the vetting context there is no "correct or "preferred" perspective to initialize the vetting model.

4. The Sample Size Profiler

In the context of the AP-Phase, after the planning parameters *Ho* & *Ha* are fixed, the next decision is:

How many observations would provide reasonable overall protocol inferential precision re: the collection of Audit Evidence?

In the context of the forecasting DSS used in the AP-Phase, the InCharge should consider the FNE and the FPE error possibilities as the inferential vetting information. It is the uniform recommendations that *a priori* one should consider balancing these two errors; tacitly this means that there should be an intelligent choice as to the sample size—this is often called the Goldilocks-accrual; not too large[The FPE-anomaly] and not too small[The FNE-anomaly], i.e., a sample size that is just-right. See Tamhane & Dunlop[T&D] (2000, p.216-7).

4.1 Sample Size: The Balancing Act for the InCharge

Here we are using the standard sample size calculation recommended, f(n:[α;β;Ho;Ha]), by T&D,p.306. This is programmed in the AP:DSS.

f(n:α;β;Ho;Ha):

$$n = ([z_\alpha \times (Ho_\% \times (1 - Ho_\%))^{0.5} + z_\beta \times (Ha_\% \times (1 - Ha_\%))^{0.5}] / Abs[Ho_\% - Ha_\%])\char`^2$$

In parametrizing f(n:α;β;Ho;Ha), the values of *Ho* and *Ha* have been noted above. We recommend that the directional *Alpha*-concern be set at 1% in the directional case. This is to be sensitive to the concerns raised by many statisticians that the Fisher-FPE concern of 5% seems to be too high for reasonable decision-making. See: Benjamin (2018) and Fisher (1925). For the *Beta*-concern, we recommend Power of 90% that is also set on the high side; anecdotally the Power is often set to 80%. However, in this vetting demonstration, there is a sample size restriction as we have collected 27 firms and for each there were five datasets collected or the illustration sample size is 135 [27 × 5] observations. This accrual was judged to be the Goldilocks-accrual. In this case, we used as the Alpha-confidence of 75% and for the Beta-concern the Power was set at 67.5%. This was the closest reasonable balance to arrive at approximately the actual accrual size.

In this context, the *Alpha*-concern is also called the Type-I Error or the False Positive action. Positive means taking an Action to reject *Ho* in favor of accepting that *Ha* is the State of Nature when in fact *Ho* [was / "could be"] the true State of Nature. Otherwise said: *Incorrectly Rejecting Ho*. A low FPE suggest a high degree of confidence [low-risk] in rejecting *Ho*. The *Beta*-concern is also called the Type-II Error or the False Negative action. Negative means NOT rejecting *Ho* when *Ha* [was / "could be"] the State of Nature. So you incorrectly believe that *Ho* is the State of Nature when in fact *Ha* is the actual State of Nature—otherwise said: *Incorrectly Failing to Reject Ho*.

In this case, the suggested initializing sample size is:

$f(n:[0.67451;0.45378;45\%;50\%]) = n$

$([(0.67451_\alpha \times (45\%_{Ho} \times (1 - 45\%_{Ho}))^{0.5} + 0.45378_\beta \times (50\%_{Ha} \times (1 - 50\%_{Ha}))^{0.5}] / Abs[45\% - 50\%])\char`\^2$

$ROUNDUP[f(n:[0.67451;0.45378;45\%;50\%]),0) = n = 127 \approx 135.$

4.2 Discussion

This is the initializing sample size for the vetting where we will examine the sensitivity of the FNE and the FPE over relevant inferential ranges. The initializing sample size is the "ball-park" sample size to achieve adequate or reasonable confidence in the decision-making context. To elucidate the core-functioning of the AP:DSS, following we will conceptually present the two core-stages of the AP:DSS profiling.

5. Core-AP:DSS Platforms: The Essentials of Vetting the AP-Test

5.1 Sensitivity of the AP[P-S]:Protocol

It would be better to "pilot-test" the proposed AP[P-S]:Protocol over various possible values—in the spirit of an Operating Characteristic Curve[OCC] sensitivity—to have some assurance that the AP protocol could likely function as expected rather than to have completed the AP-steps only to determine subsequently that there were in fact operational issues that compromised the quality of the audit evidence.

As indicated above, the inferential base of any screening test is the principal montage or the driver of the rationale justifying the use that particular protocol. Thus by extension, vetting would do well to focus on the nature of the inferential errors that "could" result from the modeling protocol—i.e., sensitivity. These are, of course the FNE usually with respect to *Ho* and the FPE usually associated with *Ha*. Therefore, the ***core-information*** upon which the vetting decision will be based is: ***The FNE & FPE Profiles for the A-priori and the Benchmarking Analyses.*** Following we will detail <u>all</u> of the steps used to generate this core-info-set and then summarize the same in a vetting profile sensitivity table.

5.2 Error Profile of the AP[P-S]:Protocol

Recall under the condition of *Ho* the client's ICoFR is assumed, for inferential testing purposes, to be the toggle-point for the EPI decision relative to the client's ICoFR]. In this regard, there are two vetting conditions the ensemble of which, will be used by the InCharge to evaluate the particular AP[P-S]:Protocol: (i) *A-priori* FNE profile, and (ii) Empirically developed FPE using a benchmarked dataset.

5.2.1 The A-priori

Vetting Assume that for the *a-priori* phase that the InCharge is interested in the FNE-profile. Assume that the Audit Team lead by the InCharge fixes *Ho* as: If 45% of the client YE-data is captured by the forecasting protocol, then the inferential decision is that the client's ICoFR is not adequate.

The rejection of *Ho* is set by the InCharge as: IF the percentage of the capture rate for AP[P-S]:Protocol is at least 50%, THEN *Ho*: [ICoFR is SEC not adequate] will be rejected and the client's ICoFR will be vetted as likely effective.

To inspect the FNE sensitivities of this test-vetting—i.e. Failing to Reject *Ho* [Accepting *Ho*] when it is NOT likely to be the assumed State of Nature—and so assuming that ICoFR of the Client's AIS protocol is not effective/adequate when, in fact, it is effective, the InCharge will use the following testing form (Note 8) suggested by T&D (2000, p.214):

FNE:OCC[μ]: P[Test Fails to Reject *Ho* Given a particular value for μ as *Ha*]

$FNE:OCC[μ] \equiv \varPhi[t \leq \frac{(50\% - μ)}{Se}]$ EQ1

Where: The *Se* is estimated using: $\hat{\pi} = Mean[(45\%: 50\%)$ & n=135 is the number of forecasting tests from the Benchmarking Accrual. This gives the *Se* of 4.3%.

A reasonable FNE sensitivity interval is: [45% to 60%]: Step = 5%.

The range specific sensitivity output of the FNE:OCC[μ] is:

Table 1. FNE Sensitivity for *Ha= μ*

μ	45%	50%	55%	60%
FNE	87.8%	50%	**12.2%**	1.0%
Risk	High	High	**Accept**	Low

Recalling: NO <u>actual</u> audit data of any kind has been produced. The FNE:OCC is a *What-If* presumption sensitivity generator; this is not a pejorative contextualization. *A priori* the InCharge will use Table 1 to make the decision using "professional" judgment. The decision context is to examine the Sensitivity—the relative change of the vetting screen of EQ1. In this case, at the 50% test value rejecting *Ho* is a "coin flip"—i.e., where *Ho* is likely to be scored as a likely reject or a likely accept with equal probability; at the next evaluation, for μ = 55% [**Bolded**], the FNE drops rapidly, the slope of the OCC is sharply negative, to 12.2%; this is certainly an acceptable FN-Risk. *Summary*: If 55% is the State of Nature, then the chance of drawing a random sampling and failing to reject *Ho* would happen 12.2% of the time or rarely. This is the FN-risk of believing that there is a need to effect an EPI when in fact this not likely to be the case given that the State of Nature was 55%.

For example, at μ = 50%, the OCC value is 0. Thus using the T. calibrated [LHS] for an exact test we have:

$\Phi[t \leq 0]$ or T.Dist[0,10000,TRUE] = 50%

For $\Phi\left[t \leq \frac{(50\% - 55\%)}{4.3\%}\right] = \Phi[t \leq -1.16335]$ = T.Dist[-1.16335,10000,TRUE] = 12.2%

Also, as the OCC presents FNE-risks thus [1-the OCC values] that are the Power of the AP[P-S]:Protocol. At 55% the test screen has Power of 87.8%; whereas at 45% the Power is only about 12%. Therefore, for the *a-priori* FNE phase, the InCharge would probably decide that the chance that the AP[P-S]:Protocol will fail to reject *Ho* when *Ha* is the minimum sensitive screening case is acceptable low. This means that if the State of Nature were to be 55% the chance of failing to reject *Ho* and believing that the client's ICoFR was in fact not adequate in the SEC context would only happen around 12% of the time. This is certainly a sensitivity profile that would engender confidence in the parametrized AP[P-S]:Protocol to be effective.

This is the first or *a-priori* FN-Phase; the next test point is the Benchmarking Phase.

5.2.2 The Benchmarking Dataset: Dipping into the Big Data Pool

The concept of using an independent testing frame for the FP-Risk offers an enriched inferential context for the AP[P-S]:Protocol-vetting. Effectively, this follows the usual logic of benchmarking the FP-risk using the FN-risk of the particular design by selecting an alternative *test-against-value*, *Ha*, for the FP-directional test to have two inferential indications: A FPE and its FNE-benchmark. To enhance the inferential impact of the AP-DSS, we have enhanced the quality of the inferential information by creating the FP-Risk using an independent empirical design. Therefore, the FP-Risk will be a relatively independent benchmarking-frame with respect to the FN-Risk. Specifically, we recommend that the comparison dataset used to benchmark the FP-Risk dimension of the AP[P-S]:Protocol be "reasonably" free from major issues that would compromise the ICoFR of the accrual set. In this case, we recommend taking advantage of the vast amount of readily available firm information available from accrual sources such as AuditAnalytics™, WRDS™[CRSP™] or The Bloomberg™ market navigation platform.

We selected randomly selected 32 firms from the Big Data set of US trading exchanges that were indexed by the Bloomberg terminals. The accrual screen used[and recommended] was: [Context: An absence of less than a few instances over the last five years] of the following: (i) required or (ii) voluntary SEC re-fillings, (iii) 302-Defficencies, (iv) 302-Weaknesses, (v) denial/temporal suspension of listing by the trading exchange(s), (vi) Non-Scope qualifications of the financial opinion, or (vii) an not adequate designation of the COSO opinion. Using the AuditAnalytics™[WRDS™] platform to collect most of this information on issues that may compromise the quality of the benchmarking as noted above, there were five firms over the Panel that had re-submissions, or 302-Deficiencies noted; thus they were eliminated. This produced 27 firms.

For each firm, we collected the following information over the Panel from the Bloomberg Terminals:

Revenue [SALES_REV_TURN(Note 9)],

[GROSS_MARGIN], [BS_CUR_ASSET_REPORT],

[BS_CUR_LIAB] &

The Current Ratio:[BS_CUR_ASSET_REPORT / BS_CUR_LIAB].

Therefore, 27 PCAOB Firms, note in the Appendix, and five variables used in the MAM model. This given 135 Reported Forecast values [27 $\times$ 5]; this is in conformity with the accrual estimated sample size.

In summary, this is the Benchmark dataset readily available from in the Big-Data world. The InCharge will now examine this dataset for the percentage of time that the YE-value is IN the 95%CIs. For this benchmarking dataset, the InCharge will NOT have the actual YE-value for the firms as this dataset usually will be collected during the Planning phase of the audit. In this case, we recommend that the InCharge holdback the last reported data point and then use that point as the surrogate for the YE-value of the previous year which, of course, is the time-indexed reality. This is a standard practice in forecasting evaluation studies. See Makridakis et al. (1982). As this is the benchmarking dataset, these empirical results will be used in conjunction with the *A-priori* OCC results as presented above to complete the two-stage vetting.

The next decision is to use the selected forecasting model, in this case the MAM[2] Excel version, to determine the percentage of YE-Value for the benchmarking dataset and then to test this against *Ha*. This will be the FPE Empirical check for a dataset where by accrual there are likely few serious AIS issues that would compromise the Client AIS Panel comparison. As we will be using the MAM[2], it would be instructive to detail the MAM[2] model to elucidate the vetting *A-priori* and *Benchmarking* component parts of the vetting protocol, which are the FPE & FNE[&Power].

The MAM is one of the standard pre-cursors to most of the various forms of the Exponential Smoothing Time-Series[TS] models. Box & Jenkins (1976). The MAM modeling Set is TS-data driven and has very few assumptions that form the calculation base. To that extent, it is an independent form of a TS-projection model. The MAM is used to create forecasts in forming the 95%CI that will be used in the AP[P-S]:Protocol.

It will be instructive at this point to use an actual dataset to illustrate the computations of the MAM model. In Table 2 are twelve Panel Points for the current ratio from 2006 to 2017] for TUI, GmbH.

Table 2. TUI, GmbH Current Ratio [CR] Panel Values

0.447317	0.534417	0.590235	0.581433	0.716273	1.052041
0.722421	0.68957	0.647043	0.616374	0.632689	0.66955

For the MAM, the Excel version of the MAM [Period=2, *Se*] was used for the dataset CR TUI in Table 2. In this case, the MAM produces the estimate of the CR. This particular MAM is a rolling two-period non-memory indexing model. The *Se* option was elected and used to form the 95%CIs in the standard manner. Using the dataset in Table 2:

The *Se* as parameterized is:

Se = 0.0143

= [SQRT(SUMXMY2 (0.6327, 0.6696; AVERAGE [0.6164, 0.6327]:AVERAGE[0.6327, 0.6696])/2)]

Using this for the Se, the MAM forecasting projection for the YE-value is 0.6511 [AVERAGE[0.6327,0.6696]. In this case, the 95%CIs for the CR are:

CR:MAM:LowerLimit: [*0.6232*]= $0.6511 - 0.0143 \times 1.96$

CR:MAM:UpperLimit: [*0.6791*] = $0.6511 + 0.0143 \times 1.96$

In the TUI case, the 2018 holdback was: [*0.5809*]. In this case, the holdback was NOT in the MAM 95%CIs.

Having illustrated the MAM Capture Intervals, consider now the dataset that was accrued for the AP-Empirical vetting stage.

The InCharge now will specify the *minimal* expected inclusion or capture rate using the MAM model in the Benchmarking context. As the MAM has a relatively smaller precision compared to the OLS regression forms, this will be demonstrated in Table 5, assume that that the minimal MAM capture rate expected for firms that have adequate ICoFR is greater than 45% as noted above—this also is a toggle-point. Using the 27 firms accrued as the AP-Benchmark, the InCharge will produce the following dataset:

The percentage of YE-values **IN** the MAM95%CIs, n=135 was **54.81%** [74/135]. The FPE test of this using the Binomial correction, even though it is not needed as the Rule of 5s is satisfied, is:

FPE[μ=45%, 74, n=135] : $[74.5 - (135 \times 45\%)] / [135 \times 45\% \times 55\%]^{.5} = 2.3787$

FPE ≡ T.DIST.R[2.3787,10000] = 0.87%.

Further, as a sensitivity test we, varied the OCC-test against value of μ=45% to 60%. The results are presented in Table 3 and are scripted with the sensitivity results in the FNE context.

The context for the FPE-sensitivity profile is: Assume that 54.81% is the capture in the AP-population. The FPE for the Rejection OCC value of 45% is 0.87% meaning that one would reject that 54.81% would likely have come from a population centered at the OCC—test against value of 45% and the toggle of the EPI as noted above. Therefore, there is very strong evidence that the testing protocol using the MAM model makes the correct decision to reject 45% as the FPE is very low. This is where the FNE comes into play. Recall that in the inferential context using the FPE as the first screen should always be paired with a FNE risk. In the vetting model that we have constructed the FPE and the FNE from different inferential contexts which is why this form of the vetting is preferred to the classical vetting where the FPE and the FNE are from the same inferential context.

5.3 The Sensitivity Results: The A-priori and the Benchmarking Ensemble The recommended inference profiling to aid in the vetting of the AP[P-S]:Protocol is to form the sensitivity profile of the FNE from the *A-priori* context, the FPE from the Empirical context and finally the Power [1- FNE] using the forms of the testing forms detailed above. If, it is the case that the FNE & FPE from these two independent inference frames are both ***in reasonable inference ranges for the*** μ-test, this would be evidence that the AP[P-S]:Protocol would be useful in investigating the forecasting test that is addressed to the COSO-Substantive disposition phase. Specifically, for the AP[P-S]:Protocol parameterized above we have profiled the OCC sensitivities for the μ-range running from 45% to 60% as was done above. We then iterated the increments by 2%.

It would be instructive to examine two cases that are in the sensitivity ranges for the FNE, FPE, & Power as presented in Table 3.

Table 3. Sensitivity Profiles for the FPE, FNE, and Power

Test μAgainst	45%	47%	49%	51%	53%	55%	57%	59%	60%
FNE[*Ha*:50%]	87.8%	75.7%	59.2%	40.8%	24.3%	12.2%	5.2%	1.8%	1.0%
FPE[54.81%]	0.9%	2.8%	7.5%	16.5%	30.5%	48.3%	66.5%	81.6%	87.3%
Power	12.2%	24.3%	40.8%	59.2%	75.7%	87.8%	94.8%	98.2%	99.0%

Discussion Table 3 is the recommended joint-sensitivity profile that will be used to vet the AP[P-S]:Protocol. Specifically, it is recommended that the sensitivities be profiled over Low-Power and High FP-Risk ranges. This will offer "negative" or undesirable trade-off zones as well as acceptable ranges. In this illustrative case, it is an excellent illustration of a vetting profile that would be instructive in the audit context as a justification for the assurance opinions. To elucidate this, consider the likely use of the information set of Table 3 in the assurance context. Recognizing that there are two independent testing frames one for the FNE and one for the FPE, the InCharge could pose, for example, two queries:

Query 1: Assume that the capture μ-test for the MAM were to have been 55%. The FPE against 54.81% is about 50%, which alone is a relative high rejection risk. However, using the FNE for 55% as the benchmark for the FPE-risk, the FNE at 55% is about 12% with relatively high Power of 88% and so would likely alley any inferential concerns re: the use of the AP[P-S]:Protocol in this instance. Simply, the InCharge can accept the High FPE Risk given the low FNE risk. This vetting information will be used as part of the audit evidence regarding the actual use of the AP[P-S]: Protocol in the execution of the audit. In this case, the vetting evidence will be used in the justification of the opinion written for the COSO opinion. This is audit evidence in the ***assurance context***.

Query 2: Assume that the capture μ-test for the MAM were to have been 45%; the FPE against 54.81% is about 1%. However, the FNE at 45% is about 90% with Power of only about 10%. In this case, the InCharge may not be willing to accept the high FNE of failing to reject *Ho*, thus launching an EPI when it is not likely to be warranted. This is an ***audit-budget-control-context*** decision.

These questions are but an illustration of the many query variants of the information that is available from the sensitivity profile of Table 3. To give advice on the possibility that the vetting evaluation of the AP[P-S]:protocol, we offer that: If there are no practical cases where either the FPE <u>OR</u> the FNE are in the Interval [≈5% to ≈25%], this would be a convincing signal that the AP[P-S]:Protocol is in need of re-tooling.   In this case, the InCharge should re-visit the audit information that was used to parameterize the AP[P-S]: Protocol. Specifically, in conjunction with the audit team. the InCharge would consider the possibility of re-organizing the constituent variable's or

re-parametrizing those that were initially used. Experience suggests that this can aid the re-orientation and produce a useful AP[P-S]: Protocol. The ideal AP[P-S]: Protocol is one for which there is a sensitivity range where both the FPE & the FNE could be located in the above interval. Experience also suggests that this is a relatively rare event—but nonetheless an achievable goal.

However useful the FPE & FNE profiles may be, there are two important caveats that should be considered.

6. Caveats

There are two issues that we need to broach to complete the examination of the vetting protocol as illustrated. They are offered as caveats—in the stipulation or elaborations sense.

6.1 CaveatA: Xx OLSR Time Series Inference From the Excel Parameter Range Model

In this section we will offer advice—experiential to be sure and so antidotal— that may aid the InCharge in effectively utilizing the vetting protocol for the AP[S-P]:Protocol. In this presentation of the suggested AP Forecasting Protocol, we used only the MAM[f( )] model. It could also have been the case that other forecasting models could have been selected by the InCharge. For example, the *On the Go Stores* used the Time Series OLS two-parameter [Intercept and Slope] Linear model [OLSR]. The OLSR is offered in the *Excel*:[*DataAnalysis*] Platform is also a reasonable choice. As the OLSR is a standard forecasting model possibility, we will briefly consider the OLSR as the AP[P-S]:Protocol model of choice. This is also programmed in the AP:DSS.

As noted by Gaber & Lusk (2018), the Excel OLSR functionality forms relatively wide-covering confidence intervals compared to the random effects and the fixed effects alternative. Therefore, this model is expected to capture all but the most extreme YE-variants.   These are effectively extreme case CI-scenarios as they are produced from the crisp-end-point parameters of the 95%CI for the Intercept <u>and</u> the Slope independently rather than jointly.

It will be instructive at this point to use an actual dataset to illustrate the Excel-computations of the OLSR; consider again the TUI-CR Panel from **2006 through 2017**].

Table 4. TUI, GmbH Current Ratio Panel Values

0.447317	0.534417	0.590235	0.581433	0.716273	1.052041
0.722421	0.68957	0.647043	0.616374	0.632689	0.66955

Here we offer the following notation for the OLSR:

Left Side [LowerLimit[LL]] 95%Boundary for actual client YE-value will be:

$95\%\text{LLimit}[\text{YE}] = [\hat{\alpha} - t_{95\%,n-2} * Se_{\alpha}] + [\hat{\beta} - t_{95\%,n-2} * Se_{\beta}] \times [\text{N} +1]$

Right Side [UpperLimit[LL]] 95%Boundary for actual client YE-value will be:

$95\%\text{ULimit}[\text{YE}] = [\hat{\alpha} + t_{95\%,n-2} * Se_{\alpha}] + [\hat{\beta} + t_{95\%,n-2} * Se_{\beta}] \times [\text{N} +1]$

Where: $t_{95\%,n-2}$ is the two-tailed t-distribution value n-2 Degrees of Freedom, $\hat{\alpha}$ is the estimated regression Intercept, $\hat{\beta}$ is the estimated regression Slope, N is the Number of points for the OLS fit.

For the TUI dataset we have the following 95%CIs computed as:

$0.17873 = [0.580318 - 2.228139 \times 0.090449] + [0.011994 - 2.228139 * 0.01229] \times [12 +1]$

$1.29375 = [0.580318 + 2.228139 \times 0.090449] + [0.011994 + 2.228139 * 0.01229] \times [12 +1]$

In this case, for the CR-TUI Panel, the 2018 Holdback is IN, as expected, given the nature of the OLSR 95%CIs, the OLSR 95%CIs. This will have implications in the setting of *Ho & Ha* in the vetting protocol as the capture rate will be relatively high compared to the MAM-smoothing model. The ratio of the averages of the precision, effectively the width of the 95%CIs of: [OLS / MAM] is 3.60—i.e., the OLS 95%CI is about 3.5 times the width of the MAM95%CI [See Table 5]; the precisions are that different.

6.2 CaveatB: Combining Model Forecasts

Collopy and Armstrong (1992) offer a clear statement of the very well established benefit of combining forecasts from various models. They note, p. 1403:

> We used equal-weights combining as the primary for a comparison [re: their RBF-Expert System] because empirical research showed that this strategy is robust. Equal weights typically provides extrapolations that are at least as accurate as other strategies across a variety of conditions.

As combining is often used in practice to improve the forecasts, we will offer some advice re: the parametrization at the vetting stage. If the decision is made to use a combining protocol to form the expectation relative to *Ho & Ha*, attention must be given to the impact of a combining protocol on precision. To illustrate the necessity of understanding the relative precision issues that must underlie the execution of the AP[P-S]:Protocol, the precision profile of the MAM, the OLSR, and the Equal-Weights [EW] combined value: AVERAGE[OLSR;MAM] are presented in Table 5.

Table 5. 95%CI-Precision for the Three Forecasting Models used in the Empirical Vetting

	Mean	Median
OLS	5,670.84	325.46
MAM	1,577.30	63.41
EW-Combining	3,624.07	197.16

As discussed above, there is an ordered set of relationals for the precision profiles. Note that EW-Combining [5,670.84;1,577.30] is 3,624.07. It is the case, that both the Parametric[Welch:ANOVA] and the Non-Parametric: Wilcoxon/Kruskal-Wallis[RankSums] are <u>overall</u> significant re: central-tendency differences: p-values <0.0001 and p-value = 0.0003. In addition, the power of the parametric test is 97.7% and so the ex-post FNE is 2.2%. All of these relationships are consistent with the nature of these forecasting models and with reports in the literature. The precision profiles reinforce the relevance of the advice that the InCharge must carefully consider the nature of the forecasting model used in the vetting stages to select a relevant parameter set as the expected inclusion of the YE-value is directly related to the magnitude of the precision.

7. Conclusion & Outlook

7.1 Conclusion

It bears repeating: The AP[P-S]:Protocol detailed above is NOT the one that would be used in the execution of the audit; rather, it is the proposed protocol that would be used depending on the inferential information generated by the vetting protocol presented above. Specifically, the following six components are the elements of the vetting or pre-testing of the AP:DSS that *may lead to founding the AP[P-S]:Protocol that will be used in the assurance audit*:

1. The Initial Sample Estimate,

2. The Selection of the Forecasting Model,

3. The *A-priori* FN-risk OCC,

4. The Selection from the Big Data Milieu the Sensitive Accounts from the Screened Sample Population,

5. The FP-Risk Drawn from the Benchmarked Analysis, and

6. The Screening of the Inferential Sensitivities for the FN, FP & Power using the Lens: [≈5% to ≈25%].

If the vetting seems reasonable according to the sensitivity screen interval: [≈5% to ≈25%], ***then*** the InCharge should include the vetting results in the current working papers as due diligence evidence that the AP[P-S]:Protocol *could be* productive and so employed in the AP-tests addressed to the COSO certification. This would be convincing evidence that "PCAOB-Best Practices" were serviced in this AP-context. Noted in another way: The Pilot test of the Proposed AP[P-S]:Protocol as effected by the AP:DSS can be used to make the decision if modifications are needed in the AP[P-S]:Protocol so as to have confidence that the AP[P-S]:Protocol could provide audit evidence in the COSO testing milieu.

7.2 Outlook: The Big-Data Context: The AP:DSS as a Facilitator

To para-phrase the *Beatles*®: <u>We</u> <u>all</u> <u>live</u> *in the Big-Data [BD] world*. The implication of the BD-context for the AP-vetting model presented in this paper has been eloquently offered by Sivarajah, Kamal, Irani & Weerakkody (2017, p.263):

> *Big Data (BD), with their potential to ascertain valued **insights** for enhanced decision-making process, have recently attracted substantial interest from both academics and practitioners. Big Data Analytics (BDA) is increasingly becoming a treading practice*

that many organizations are adopting with the purpose of constructing valuable information from BD.

This *insight* is the linkage that we have used to create an embedded structure the driver of which is audit judgment and so is at the nexus of Big Data & DSS-structure & Judgmental forming of the final protocol re: the AP-inferential evaluation. This is to say that Big Data is the platform from which the analytics are formed and inform the decision-making process. This linkage is also emphasized by Gandomi & Haider (2015, p.140):

> *Big data are worthless in a vacuum. Its potential value is unlocked only when leveraged to drive decision making. To enable such evidence-based decision making, organizations need efficient processes to turn high volume of fast-moving and diverse data into meaningful insights.*

Note again the use of *insight* as the "deliverable" of the platforms that are the processing links in the Big Data world. This is exactly the reason that we have created the AP:DSS so as to leverage the fact that there are myriad numbers of client accounts in the BD-world that can be used in the benchmarking phase so as to create relevant AP-inferential information sets. This is to say: to effectively and efficiently and economically process the extensive datasets that are part of the client's AIS to create the audit evidence that justifies the COSO opinion and the Assurance re: the client's financials, the power of e-World must be used. For example, as a personal communication from Mr. Richard E. Geoffroy, CPA: Partner: M&A Tax/Private Equity KPMG:Boston during Tax-brainstorming sessions, a question often posed is: *How can we use Watson's cognitive technology?* (Note 10)

Ms. Mason, [see: Drew (2019, p.22)] gives the closing for the homily of our research report:

> "- - -, *I am tremendously excited about advanced analytics and what's happening in data science. The ability to take data—not only from historical sources, but current sources that are not only quantitative but qualitative—and to analyze them in real time, plus adding the ability to do predictive analytics on top of it, gives me this nerd excitement that I can't even describe in words".*

Yes Ms. Mason, we agree. This excitement actually was the motivation of our research; paraphrasing the recent report of Cohen, Rozario & Zhang in *The CPA Journal* (2019), it is the best of times and we are empowered to use technology such as the AP:DSS or Robotic Process Automation [RPA] to improve the effectiveness and efficiency while maintaining control of the cost of creating such DSS tools. Also relative to the forecasting context, it is important to have well-designed GUI-enhancements for the DSS see: Fildes, Goodwin & Lawrence (2006).

The takeaway implication is obvious: If you cannot figure-out how to "Partner with technically efficient Decision Support Systems" and leverage the Big-Data data sources your Audit LLP will never be engaged by firms who live in the Big-Data milieu; if by some remote chance your Audit LLP happened to be engaged, after a few cycles you would be dismissed just as Brad and Ken were summarily dismissed by *Watson*.

Acknowledgments

References

American Institute of Certified Public Accountants [AICPA]. (2012). *Audit guide: Analytical procedures*, American Institute of Certified Public Accountants, Inc. New York, NY USA: https://competency.aicpa.org/media_resources/208125-analytical-procedures-aicpa-audit-guide

Arens, A., Elder, R., Beasley, M. & Hogan, C. (2015). *Auditing and assurance services: An integrated approach*. Pearson, 16th Ed. ISBN: 0-13-406582-3

Benjamin, D. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*, 6-10. https://doi.org/10.1038/s41562-017-0189-z

Bettauer, A. (1975). Extending audit procedures: When and how. *Journal of Accountancy, 11*, 69-72.

Box, G. & Jenkins, G. (1976). *Time series analysis, forecasting and control*. Holden-Day, San Francisco, 3rd Ed. ISBN: 9780816211043

Collopy, F. & Armstrong, J. (1992). Rule-Based forecasting: Development and validation of an expert systems approach to combining time series extrapolations. *Management Science, 38*, 1394-1414. https://doi.org/10.1287/mnsc.38.10.1394

Cohen, M., Rozario, A. & Zhang, C. (2019). Exploring the use of robotic process automation (RPA) in substantive audit procedures: A case study. *The CPA Journal*, *July*, 49-53.

Drew, J. (2019). What's 'critical' for CPAs to learn in an AI-powered world. *Journal of Accountancy, June*, 20-24.

Gandomi, A. & Haider, M. (2015). Beyond the hype: Big data concepts, methods and analytics. *International Journal of Information Management, 35*, 137-144   http://dx.doi.org/10.1016/j.ijinfomgt.2014.10.007

Fildes, R., Goodwin, P. & Lawrence, M. (2006). The design features of forecasting support systems and their effectiveness. *Decision Support Systems, 42*, 351-361. https://doi.org/10.1016/j.dss.2005.01.003

Fisher, R. A. (1925). *Statistical methods for research workers*. Oliver & Boyd Press, Edinburgh: Scotland.

Fukukawa, H., Mock, T. & Wright, A. (2011). Client risk factors and audit resource allocation decisions', *ABACUS, 47*, 85-108 [online] http://dx.doi.org/10.1111/j.14676281.2011.00332.x. .

Gaber, M. & Lusk, E. (2018). Analytical procedures phase of PCAOB audits: A note of caution in selecting the forecasting model. *Applied Finance and Accounting, 4*, 76-84.   https://doi.org/10.11114/afa.v4i1.2811

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R. & Winkler, R. (1982). The accuracy of extrapolation (Time Series) methods: Results of a forecasting competition. *Journal of Forecasting, 1*, 111-153.   http://dx.doi.org/10.1002/for.3980010202

O'Donnell, E. & Perkins, J. (2011). Assessing risk with analytical procedures: Do systems-thinking tools help auditors focus on diagnostic patterns? *Auditing: A Journal of Practice & Theory, 30*, 273-283. [online] http://dx.doi.org/10.2308/ajpt-10148.

Rosenthal, R. (1991). *Meta-Analytic procedures for social research*. Sage Publishers. http://dx.doi.org/10.4135/9781412984997

Sivarajah, U., Kamal, M., Irani, Z. & Weerakkody V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research, 70*, 263-286.   http://dx.doi.org/10.1016/j.jbusres.2016.08.001

Tamhane, A & Dunlop, D. (2000). *Statistics and data analysis*. Prentice-Hall: Upper-Saddle River NJ. USA ISBN:013-744426-5

Wheeler, S. & Pany, K. (1990). Assessing the performance of analytical procedures: A best case scenario. *The Accounting Review, 65*, 555-577.

Appendix

ABI	ALL	CVX	DAL	DREL	EBAY	FDX	GIS	H
HP	IBM	INTC	JACK	JCP	Macy[M]	MET	MSFT	NCLH
PZZA	Samsung	SIX	TGT	TRV	TUI	VZ	WBA	WWW

Table A1 PCAOB Accrual for Empirical Benchmark Vetting of the AP[P-S]:Protocol

The AP:DSS Vetting Profiler Following, we give an overview of the platforms of the AP:DSS. Of note, we have integrated the AP:DSS in our AIS and Audit and Assurance courses. We are happy to share the problem/exercise sets used in these courses. Finally, the AP:DSS is available as an e-Download at no cost.

Platform 1: UserForm: *Computes the Sample Size.* **Inputs: [Ho%, Ha%, α, β]**

Platform 2: LaunchButton: *Computes: (i) The 95%CIs for OLSR & MAM & Average, (ii) The Precision, and (iii) the Inclusion Profile of the Holdbacks*. **Inputs: The Benchmarked Accrual Data**

Platform 3: LaunchButton: *Computes: Table 3 noted above:* [*FPE & FNE & Power*]. **Inputs: Platform 2 Computations and Sensitivity Calibrations**

Processing Time: Average [< 8 seconds]

Notes

Note 1. The PCAOB requires the use of AP in the execution the assurance audits for traded organizations. The following link: *https://pcaobus.org/Standards/Auditing/Pages/default.aspx* opens access to the current version of AS5. Attention, AS5[2007] has been extensively modified. The current version of AS5 is 15Dec2017. Audit Standard 2 (AS2, http://pcaobus.org/Standards/Auditing/Pages/Auditing_Standard_2.aspx) of the Public Company Accounting Oversight Board [PCAOB: Sarbanes-Oxley: Pub. L. 107204, 116 Stat/745 (2002)] which is the Public Accounting LLP licensing partner of the Security and Exchange Commission (SEC). In recent updates to AS5[15Dec2017

Note 2. From the AICPA: https://www.aicpa.org/Research/Standards/AuditAttest/DownloadableDocuments/AU-00150.pdf. Also instructive is: https://www.investopedia.com/terms/g/gaas.asp

Note 3. See: https://www.dhs.gov/news/2018/02/06/national-vetting-center. Vetting is becoming very important as a pre-analysis data-conditioning phase.

Note 4. We also had the benefit of the council provided by Dr. Manuel Bern, The Chief of Internal Audit of *TUI*, *GmbH*. Dr. Bern was previously a forensic audit specialist with Deloitte Touche, Frankfort, Germany.

Note 5. The PCAOB[SeeFn[i]] notes in the Documentation of Specific Matters[.10:p.51-52] section: Documentation of auditing procedures that involve the inspection of documents or confirmation, including tests of details, tests of operating effectiveness of controls, and walkthroughs, should include identification of the items inspected. Documentation of auditing procedures related to the inspection of significant contracts or agreements should include abstracts or copies of the documents. Further important details are found in sectional Paragraphs[37.p 112 &38.p 113].

Note 6. https://www.sec.gov/rules/final/33-8124.htm

Note 7. *Here **assumed** to be: Excel™[DataAnalsis[Moving Average Model[Period2]]].*

Note 8. This T&D form may seem to be a variant of the standard forms of the FNE functions. However, if the InCharge would have fixed the *Ho* rejection at the α-rejection point then this would be identical to/with the two standard forms:

T.DIST(t − (ABS(Ho − Ha)/Se), Df, True) ≡ T.DIST((FPE-Value − Ha)/Se), Df, True). For example, at a 95%FPE rejection the cut-point would be: 52.07%. In this case, the T&D z-value would be -0.6817. This is the same value as [1.645006 − 2.326700] and [(52.07% - 55%) / 4.3 = -0.6817].

Note 9. These are the actual labels used by the Bloomberg Platform at the time we accrued the Benchmark dataset.

Note 10. *Watson* is the IBM AI-Decision-making platform that is becoming ubiquitous in the decision-making world. A demarcation in human history is the fact that *Watson* is currently the reigning Jeopardy™ Champion winning The head to head to head computation: {Watson v. Brad Rudder v. Ken Jennings}. See: https://www.techrepublic.com/article/ibm-watson-the-inside-story-of-how-the-jeopardy-winning-supercomputer-was-born-and-what-it-wants-to-do-next/