Received: 28 January 2017 | Revised: 18 August 2017 | Accepted: 22 August 2017
DOI: 10.1111/ijau.12103

WILEY

ORIGINAL ARTICLE

Managing group audit risk in a multicomponent audit setting

Lynford Graham[1] | Jean C. Bedard[1] | Saurav Dutta[2,3]

[1] Bentley University, Waltham, MA, USA
[2] Curtin University, Bentley, Western Australia
[3] University at Albany, State University of New York, Albany, NY, USA

Correspondence
Lynford Graham, Bentley University, Waltham, MA, USA, c/o 298 Taylor Rd S Short Hills, NJ 07078, USA.
Email: lgrahamcpa@verizon.net

This paper considers the challenges in planning the scope of auditing procedures in a group audit setting for an entity with geographically dispersed components which vary in risk characteristics. Auditing all the components for a complex group entity is often infeasible, hence the auditor faces risk from components not audited, as well as the normal sampling risk resulting from applying audit procedures to certain components. Auditing standards do not explain how to consider the risk factors and consider what portion of a multiple component entity should be selected for audit to be able to issue an unqualified audit opinion on the group. In this paper we describe a step-by-step method for determining a minimum number of component audits needed to support an aggregate low level of audit risk of material misstatement. The paper responds to calls from academics, practitioners, and standards-setters for theoretically valid and practically feasible solutions to the group audit problem, using a method that combines professional judgment and experience with basic statistical principles in an ensemble approach.

KEYWORDS
audit methodology, audit planning, sampling

1 | INTRODUCTION

The recent focus on the inherent complexities of conducting the audit of a large, dispersed entity in a group audit setting (e.g., Public Company Accounting Oversight Board [PCAOB], 2016) has brought to the forefront a longstanding issue in audit sampling: how to control for the risks associated with performing audit procedures at a limited number of components or locations in a multiple component entity. Since 1938, when the inventory misstatement at McKesson and Robbins was revealed (e.g., Dutta, 2013, p. 62), auditors have been challenged to provide reasonable assurance on the fair statement of aggregate financial statements when components exist at multiple, geographically dispersed locations, while simultaneously controlling audit costs. The extensive use of component auditors while conducting a global audit accentuates the challenges with sampling, materiality, audit effort allocation, and evidence aggregation. Notably, reports from the PCAOB and other auditing regulators continue to highlight "group audits" as a source of concern.[1] While this issue is commonly discussed in the context of global group audits for Big 4 firms (e.g., Downey & Bedard, 2017), it is also a real and practical audit issue for an increasing number of medium and smaller entities that have a profile for which allocating audit effort across components or subsidiaries is needed.

The purpose of this paper is to extend the professional and academic literature by providing a justifiable and flexible method for determining a minimum number of components (or locations) to audit and the assurance needed at those selected units, in order to provide a desired level of overall audit assurance. This method is most applicable when a large number of components or locations exist and it is impractical to audit (or visit) all significant (individual or in the aggregate) components.

Current professional guidance for group (multinational) audits—ISA 600 (International Audit and Attest Standards Board, 2009) and AU-C section 600 (American Institute of Certified Public Accountants [AICPA], 2012d)—require allocating overall audit effort to components,[2] but does not provide a methodology for doing so. Consequently, implementation problems have arisen and auditors are having difficulty with demonstrating compliance with the standard as noted in endnote 1. Typical audit sampling methods also do not provide guidance to the auditor on controlling the risks while minimizing cost, when entity operations are geographically dispersed and preliminary audit planning reveals that components vary in characteristics suggesting a differential risk profile (e.g., size of component, internal controls effectiveness). Extant methods in the academic literature for allocating audit effort to elements or components of the auditee (Dutta & Graham, 1998; Elliott & Rogers, 1972; Glover, Prawitt, Liljegren, & Messier, 2008; Stewart & Kinney, 2013) do not directly address the risk associated with unsampled components. These methods assume that all elements that are significant, individually and in the aggregate, will be audited. However, as recently noted in a summary of practice issues in multinational auditing (Sunderland & Trompeter, 2017, p. 170), group audit leaders must be aware of the potential impact of nonsignificant components:

> *It is very often the case in a multinational environment that there are very few quantitatively significant components. It may be the case that the residual balances do not represent 20% of the consolidated amounts; rather they represent 60% of the balance. A question that is always difficult to address is how this residual population should be thought of in the context of the audit.*

Without a method for dealing with the residual balances beyond those considered significant, material errors may go undetected and the auditor exposed to unforeseen risks.

In this paper, we address this important practice issue by describing a step-by-step method for determining a minimum number of components (or locations) to audit and the assurance needed at those selected units, which will yield the desired level of overall audit assurance. In applying this or any other approach in the multiple component context, the auditor must first apply judgment to identify the components within a group.[3] This can be a difficult problem in practice. While some similar components may be treated as a unit, client characteristics may prevent aggregation for others. For instance, when examining original source documents and observing physical assets such as inventory and property, and plant and equipment, physical visits may be required at the location where the assets reside. Depending on the centralization of the accounting records and access to source documents, auditing some components or accounts within components requires the presence of the auditor, while others may be tested or audited remotely.

While the approach described in this paper is based on statistical principles, it can be adapted for application in a judgment-oriented environment. In addition to a series of structured examples illustrating the method, we present a generalized model of the approach, and discuss possible extensions for developing more robust approaches that explicitly consider more of the potential qualitative factors. This more structured methodology will be useful in helping firms meet expectations in PCAOB and AICPA auditing standards for documenting the process undertaken and results of the required risk assessment, as well as the rationale for testing decisions. These include the decision about which components will be subject to substantive tests of detail and other audit procedures, and which components may only be subject to analytical procedures.

The remainder of the paper proceeds as follows. In Section 2 we describe professional guidance regarding the component sampling problem and discuss the existing approaches to the problem. Section 3 describes a structured method to allocate audit effort to components, and illustrative numerical examples are presented in Section 4. We conclude in Section 5 by considering implications and extensions of the proposed method.

2 | PRACTICE GUIDANCE AND PRIOR RESEARCH

2.1 | Developments in professional practice

The problems of auditing businesses with financial statement elements distributed across diverse components or locations have long been evident to audit practice and regulators. Periodically, startling revelations of intricate fraud schemes rattle the accounting profession, such as Allied Crude in the 1970s, Phar-Mor in the 1990s, and Rite-Aid in the early 2000s (Securities and Exchange Commission, 2002).[4] These auditing challenges have been exacerbated in the global economy, and the concern is not limited to the USA. The International Audit and Attest Standards Board (2009) sought to address this longstanding problem by promulgating ISA 600, which requires auditors to adopt both administrative and performance requirements when an audited entity consists of multiple identified components. It requires auditors to define component materiality and requires that component materiality be less than group materiality. This requirement was instituted to explicitly allocate audit effort to components, so that when combining the results of tests of accounts and balances at the component level the aggregation could result in a conclusion of a low aggregate risk of a material misstatement at the group level.[5]

In the USA, the Panel on Audit Effectiveness (2000, pp. 21–23) recognized the difficulties in auditing dispersed entities with operations in multiple locations (components), and recommended that the US Auditing Standards Board develop guidance.[6] In conjunction with the Clarity "suite" of standards released by the AICPA in 2012, ISA 600 was adopted with minor changes into US Generally Accepted Auditing Standards (AICPA, 2012c). This issue has also been considered by task forces responsible for developing the AICPA's audit guide *Audit Sampling*. Awareness of questions arising from audit practice about how to demonstrate compliance with the "low audit risk" opinion requirement of existing standards led to inclusion in the 2008 and subsequent editions (e.g., AICPA, 2014, appendix E) of an appendix describing a two-stage risk associated with audits with multiple components. The adoption of this approach notes consideration of risk in two stages: (i) selection risk (the risk associated with not selecting some components for audit procedures); and (ii) detection risk (the risk of not performing sufficient procedures at a component or location to identify a misstatement condition of importance).[7] However, it does not provide any application guidance or examples to implement the concept. In 2013, the AICPA issued a risk alert regarding group audits (AICPA, 2013). While the alert provides a number of qualitative and quantitative factors to consider regarding selecting components or locations, it contains no quantitative guidance on selecting components or setting the scope of audit procedures at specific components. The recently published *Audit Guide: Assessing and Responding to Audit Risk in a Financial Statement Audit* (AICPA, 2016) includes an appendix summarizing some guidance from the *Group Audits* alert, and clarifying the wider applicability of the aforementioned AICPA *Group Audits* standard. The post-Clarity versions of the audit guides on *Audit Sampling* (AICPA, 2014, table 4–3) and *Assessing and Responding to Audit Risk in a Financial Statement Audit* (AICPA, 2016, appendix J) present a table to assist auditors in assessing the factors to be considered in providing an appropriate reduction of materiality for setting measures such as performance materiality[8] and tolerable misstatement (a threshold value in designing and evaluating sampling and other detailed tests). However, such guidance does not address the complex issues of identifying components or how many components should be audited.

2.2 | Current audit practice

In the absence of detailed guidance from professional standards, many audit firms have historically relied on methods such as "dollar coverage" to gain assurance regarding the fairness of the aggregate financial statements. When a relatively few components comprise a large portion of the entity, such methods can be effective. However, expansion and globalization of businesses have caused methods based on dollar coverage to be less effective as a general strategy, since in many cases with larger and complex entities the concentration of assets, revenues, or other bases of allocation may be widely dispersed and a large portion of the entity cannot be selected by identifying a reasonable subset of components. Thus, the risk of misstatement associated with components not audited increases, and should not be ignored.

A recent survey of sampling policies of US auditing firms (Christensen, Elder, & Glover, 2015) does not specifically identify an approach to determine the number of components to be selected, but instead identifies several firm policies addressing sample sizes to be used at the selected components. How the firms identified the components where procedures would be applied was not addressed in any response. Christensen et al. (2015) conclude that the sampling policies differ significantly between the larger firms.[9]

We have observed that, in practice, such scoping decisions often involve selection of the largest components and a few others based on "judgment." The use of "judgment" is often cited in describing how the scope of component selection was determined. In some cases, specific risks of misstatement are considered in this selection. However, such unstructured scoping decisions are unlikely to be able to clearly demonstrate meeting the requirements for allocation of audit effort in the group audit standards. In some recent updates, the policies of some firms have simply cited the basic requirements of the standard: that component performance materiality be set at less than component materiality and that component materiality be less than group materiality. However, we are aware of a few firms (large and small) that have experimented with implementing the two-stage multilocation risk formulation as described in the AICPA guide *Audit Sampling*.

2.3 | Review of relevant literature

Academic research providing guidance on audit materiality and auditing components or financial statement elements has adopted one of two approaches: quantitative, leading to normative prescription; and qualitative approaches that identify factors for consideration or are descriptive of actual practice. In the former track, a seminal discussion linking statistical concepts to the allocation of effort was developed by Elliott and Rogers (1972). In that article, a method was developed to combine (aggregate) the statistical limits of independent test results from various financial statement accounts. This approach provided a way, borrowing from statistical sampling principles, to set the scope of work in the accounts so that the results could aggregate to form a conclusion that there exists a low risk of aggregate material misstatement. A subsequent study centered on the use of classical statistical sampling applied to combining the results of samples from different populations (Kim, Neter, & Godfrey, 1987).

Dutta and Graham (1998) extended the concept of quantitatively allocating audit effort amongst different accounts in a set of financial statements by including a cost factor. Their model, which also rests on classical statistical theory, further allows for varying levels of tolerable misstatement in some accounts, while achieving a low overall level of audit risk. Glover et al. (2008) illustrate several methods of allocating materiality to significant components, specifically suggesting unaudited components can be examined using analytical procedures. More recently, in a model specific to the group audit context, Stewart and Kinney (2013) use a Bayesian statistical approach based on the gamma distribution to allocate materiality to components. All of the aforementioned methods explicitly or implicitly assume that the components or elements not included in the allocation are not significant.

In the qualitative stream of audit materiality research, Allen and Loebbecke (1998) combine a literature review with interviews of several experienced auditors and a review of a few certified public accounting firm policy manuals[10] to enumerate the various factors employed in scoping component audits during planning. While some factors they identify are quantitative (e.g., number of locations, size of locations), others are qualitative in nature. Subsequent to the issuance of the group audit standard, some recent papers address the issue of allocating audit effort to components. Westervelt (2014) discusses issues to be considered when assessing if the group audit standard applies, and factors that may be important when allocating audit effort among components in a group audit setting. Additionally, Sanders (2014) proposes a list of qualitative factors to consider when selecting specific components for audit. Some of these considerations could be incorporated as extensions of the formulation in this paper. However, the identification of components itself remains complex and subjective.

In 2011, the PCAOB reached out to academia for syntheses of research on key topics most relevant to audit quality, including "multilocation" audit issues. Responding to that call, Asare et al. (2013, p. 149) state that "there is currently no generally accepted scoping model" for multilocation (or multinational) audits, and call for research "on model development as well as how materiality and risk relate to the multi-location audit environment."

This paper responds to that call by developing a structured approach to quantify and control the risk of material misstatement in group audits where it is not feasible or practical to subject all components to audit procedures beyond analytical procedures. Further clarifying this problem with details from current audit practice, Sunderland and Trompeter (2017) note that, after identifying significant components for examination, the unexamined "residual" may be a very large proportion of aggregate value (e.g., 60% of the balance). They ask how the auditor should address the risk associated with these unsampled components. As noted by Asare et al. (2013), neither the qualitative nor quantitative research literature to date provides a mechanism to help auditors plan the scope of their procedures in the audit situation where it is impractical to audit all but an insignificant portion of the group entity, as is commonly encountered today in large, complex engagements.[11] Existing models and approaches do not specifically address the risk that arises when some significant (alone or in the aggregate) components will not be tested, and do not illustrate any allocation of materiality to the components that are proposed to be examined only through analytical procedures and thus do not consider any risks associated with unexamined components. The formulation in this paper considers such risks. Qualitative methods acknowledge the problem, but to date provide no structured solution.

In practice, auditors' approaches have been based on the premise that when sufficient evidence regarding a large portion of the dollars has been obtained, the risks associated with items at unexamined components or locations would be mitigated. While this may suffice in some circumstances, in the current business environment it is becoming increasingly impractical. Retail drug stores, licensed operations, and clothing chains have hundreds of outlets of rather similar size, precluding the efficient examination of the bulk of the assets or revenues or other measures when applying audit procedures to only a few components. Additionally, large multinationals may have operations in diverse geographical locations. Setting the scope of components to be audited is especially important when demonstrating low audit risk in an environment where a large number of audit components exist, and when data and supporting evidence are not accessible in a central location. In many entities with many components, evidence supporting the component financial statements needs to be gathered at the component site. Sunderland and Trompeter (2017, p. 170) provide an example from a PCAOB inspection critiquing a firm's failure to address risks at various locations:

> *One issuer had numerous foreign locations, which accounted for over 20 percent of the issuer's revenue. The Firm did not visit any of the foreign locations in connection with the audit, nor did it use the work of its international affiliates or other auditors in reporting on the issuer's financial statements in the year under audit. The issuer's foreign locations did not have common information technology systems, processes, or controls. (PCAOB Inspection report, inspection findings, annually inspected multinational audit firm).*

If the auditor infrequently or never directly obtains evidence from certain locations, or predictively audits only the same locations over time, risks of not identifying a pattern of misstatement rise (AICPA, 2012a). The greater the aggregate materiality of items in various components that may not be subjected to audit procedures, the higher the risk that discrepancies and misstatements can go undetected for long periods. Analytical procedures may be insufficient to detect consistently incorrect (either overstated or understated) results from the unexamined components, as there may not be adequate audit evidence of what the results should be (the benchmark, or expectations) from which to identify anomalies. In addition to verifying the existence and valuation of component financial statement amounts, audit procedures can provide evidence regarding the completeness of liabilities and may clarify the implications of transactions with related parties.[12]

While the importance of obtaining sufficient audit evidence at components of entities has been underscored by ISA 600 and subsequent guidance from national regulatory bodies, a formal approach to help determine a minimum number of components to be audited and quantification of the risk emanating from nontesting of some components has not been proposed in the research literature. This paper addresses three primary concerns. First, how can one determine a minimum number of components that can be reliably tested in order to provide a planned level of assurance? Second, what scope (extent) of procedures is required at the selected components to provide sufficient evidence supporting or questioning whether material misstatement might exist in the aggregate entity? Third, how might certain qualitative factors that contribute to or mitigate risk be considered when selecting components to audit?

3 | A FORMULATION THAT CONSIDERS RISK FROM COMPONENTS NOT AUDITED

The methodology illustrated in this paper combines professional judgment and experience with fundamental statistical practice in an ensemble approach[13] that may be more effective than either approach alone. Prior to determining the minimum number of components to be audited, the auditor will usually perform several steps, such as: risk assessment; determining an appropriate surrogate for selection risk, such as asset size, revenues, or profit contribution; setting desired assurance level targets; and considering possible patterns of misstatement. We next discuss these preliminary steps.

3.1 | Preliminary steps

3.1.1 | Assessing risk

The audit risk model in the current professional literature (AICPA, 2014; PCAOB, 2010) and as historically followed in practice (Lemon, Tatum, & Turley, 2000, p. 17) describes the elements of inherent, control, analytical procedures risk and substantive tests of detail and how they relate to overall audit risk. These elements would normally be assessed for all relevant assertions and accounts for all components. After considering these risks, the remaining (low) overall audit risk can be estimated.

The results of the various assessments can help specify which components would be more effective to audit due to size, risk, or other characteristics. When the relative size of various components and entity-wide assessments of internal control and other factors reveal some degree of homogeneity in those factors across the entity, the audit problem is somewhat simplified.

3.1.2 | Identifying individually large, high-risk and insignificant sites

A critical input when setting the scope of planned procedures is to determine a group materiality threshold. Materiality is a judgment, as noted in both the auditing and the accounting literature. Component materiality is a lesser amount and may be determined after considering the factors identified in the aforementioned AICPA audit guides. Once component materiality is established for the identified components, the auditor can better assess how the relative size of the components can be used in the selection of which components are to be audited. When the entity comprises only a few significant components, the allocation of audit effort to those components is a simpler task, and one that seems adequately addressed by the existing literature. When the number of components is very large and cannot be simplified by just selecting the larger components, the selection of which components to select for applying audit procedures is more complex, because there is a risk that needs to be recognized when not applying audit procedures to all of the components.

As a preliminary step in addressing this issue of many components, components that are very large or believed to be highly prone to misstatement can be segregated out and targeted for definite audit attention, because the risk of undiscovered misstatement in those components may be too high if not audited. This identification of unusual components also contributes to more homogeneity in the remaining components and reduces the risk that, if such an unusual component is selected in a sample of components, the results of audit procedures may extrapolate to give a false indicator of the condition than actually exists in the true population. Also, components that even in the aggregate could never be of consequence to the audit opinion can often be identified and segregated from the population of components. Barring concerns that the recorded values might be vastly understated, sites of clearly no audit consequence may be set aside through considering their inherent risk assessment, internal controls assessment, and then applying analytical procedures. Which audit procedures, beyond analytical procedures, are applicable to these components is a matter of judgment.

Another possibility for simplifying the population of components can be to separate out any component entities where a full-scope audit resulting in an audit opinion is to be conducted due to regulatory, statutory, or legal requirements. The criteria often associated with a full-scope audit of a component would ordinarily be more conservative (lower) than the scope of any allocated materiality from the group audit perspective. By further reducing the population of potential components, the overall audit effort may often be reduced. For the components that remain after this process, the formulation in this paper may be used to support a minimum number of components to be examined from the remaining pool of components, subsidiaries, or locations. The auditor may use random or haphazard (without conscious bias) selection criteria in the absence of more risk-sensitive selection procedures, as discussed later in this paper.

3.1.3 | Targeting an acceptable level of overall risk

The determination of the overall desired assurance or risk at the group level is also an auditor judgment. While the basic audit risk model in the AICPA *Audit Sampling* guide illustrates a target audit risk of 5%, there is no professional requirement to use that specific risk level. However, higher risk levels may call into question whether professional standards in achieving reasonable audit assurance have been followed. Once the group level of materiality is determined, that target is used to assess performance materiality at the group level and then the component materiality at the component (or location) level. That sets in motion the planning of procedures to be performed at the component level so that the component results aggregate and can support issuing an opinion at an appropriate low level of audit risk.

Setting selection and detection risk

To control risks of aggregate misstatement in a multiple component situation, the auditor should first control the risk that too few components will be audited to reliably detect a plausible pattern of misstatement that might aggregate to a material misstatement. Secondly, the auditor needs to control a second risk (detection risk) -- that the procedures applied at each component may be insufficient to reliably detect a plausible misstatement pattern at that audit unit.

In a simple quantitative sense, one can multiply the *complement* of these risks (1.0 − selection risk, and 1.0 − detection risk) to approximate the overall assurance from the test.[14] As an example of this concept, where the auditor has achieved a fairly high assurance (e.g., a risk of no more than 10% for each risk) on *both* selection risk and detection risk, the overall assurance would be at least 81%:

$$\begin{aligned}\text{Overall assurance} &= \overbrace{(1.0\text{–}0.10\text{ risk})}^{\text{Selection risk}} \times \overbrace{(1.0\text{–}0.10\text{ risk})}^{\text{Detection risk}} \\ &= 0.81, \text{ or } 81\%\end{aligned}$$

This calculation suggests that the value of detection risk is statistically independent of the value of selection risk. By using a risk-based structured model, the auditor should be able to design a better audit than by simply randomly selecting a few components and applying unspecified effort at each unit selected. Historically, the method used in practice to determine how many components and which components would be subject to detailed audit procedures has been documented as a "judgment." Understanding the relationship between risk factors at each step is a way in which professional judgment and experience can combine with fundamental statistical practice in an ensemble approach that may be more effective than either approach alone.

Note that the combined assurance is always less than the smaller of the individual assurance levels of the tests. Even achieving moderately high assurances (e.g., 90%) for each of the two stages can result in a more moderate overall assurance (e.g., 81%) when both stages are considered. This property of two-stage sampling may not always be obvious to practitioners when applying only judgment. To simplify the explanation, if one could apply effective audit procedures to the financial statements at each audit component, this should reduce the detection risk, making the selection risk the more important risk to control (AICPA, 2014, appendix E).

Selection/detection risk and assurance can also be rebalanced to provide the overall assurance desired, to minimize or at least control the individual fixed and variable audit costs that determine the economics of the audit. If the fixed costs of auditing some or all of the components are relatively high, a plan to audit the fewest number of sites and planning for a low detection risk at the audited units will be best. Formulating the cost versus risk trade-offs for each audit component is an important extension for future research and is noted further in the model illustrated in Appendix A.

To estimate a targeted selection risk for applying the proposed formulation, some important factors to be considered are:

- The results of inherent risk and control risk assessments as they relate to the components or locations not yet selected for more intensive audit procedures.
- The nature and extent of other substantive procedures that contribute to the audit objective. For example, if regulators or internal

audit have performed procedures to address various audit risks at the component or location level, this can be used reduce the level of independent auditor procedures at that selected unit.

For example, when entity-wide internal controls over financial reporting are assessed as effective, the assurance from substantive and other procedures at the various components may be reduced accordingly. If internal audit procedures are reliable and objective and contribute to the auditor's assurance concerning the components, then the assurance required from testing at the components may be reduced.

Thus, rather than seeking “high” overall selection and detection assurance in the selection of components, a more “moderate” level of assurance may be justified when other sources of evidence such as inherent and control risks at the components reduce the need for substantive test assurance. The implication is that these conditions support lowering the number of components to be audited due to the reliance on these other factors. Table 1 shows the impact that some factors might have on the assurance parameter for selection assurance.

While Table 1 is presented as a sequence of dichotomous choices, practitioners should recognize that these are just the end points of a continuous scale. A fundamental statistical property of risk factors is that they can be expressed as probabilities, and thus, as in the aforementioned selection and detection risk example, can be combined. The mathematical implication is that the resulting overall risk (combined selection and detection risk) cannot be lower than either the selection or detection risk. A high selection risk cannot be reduced by setting a lower detection risk at the audited components; however, a balance between the two risks can be made to consider costs.

3.2 | Determination of a surrogate measure for selection

Before selection of components can be made, a surrogate measure needs to be identified as a basis for the selection.[15] Glover et al. (2008) illustrate the use of revenues or assets as a potential base for the allocation of audit effort between components when using their formulation. In addition to these measures, other measures could be used to make the allocation, such as profit contribution or net assets. In advance of a more refined method than that proposed in this article, a measure that correlates best to the perceived risks in the pool of components may make the most sense. In the illustrations that follow, it is assumed that total assets are the most appropriate surrogate for selection risk.

3.3 | Considering possible patterns of misstatement

To design a test that can result in a minimum number of components or locations to be audited, the formulation in this paper requires some possible or plausible patterns of misstatement to be considered.[16] From a specified pattern of possible misstatement, the number of components or locations can be determined by finding how many of these need to be selected to ensure that, at a specified level of selection risk, at least once example will appear in the sample, indicating that such a misstatement pattern might exist in the population.

For example, if group materiality (after removing separately the large, risky, full-scope audit and insignificant sites) is 10% of the remaining component asset balances, then one pattern of possible misstatement that could produce material error is that all components could be misstated by exactly 10%. If this were the case, auditing one component and auditing to within this tolerance and finding such a misstatement would be sufficient to signal that further tests are needed. However, the auditor cannot know in advance whether this pattern exists, and such a pattern would likely be rare. More commonly, misstatements would be differentially distributed across the various components. The procedure described in the following is designed to detect the more difficult (i.e., “needle in the haystack”) error pattern—where a small number of components are totally or highly misstated, but could aggregate to a significant misstatement.

4 | ILLUSTRATION OF THE METHOD

In this section we illustrate the method through three numerical examples. The first example makes the simplistic assumptions that: (i) all the components are of the same size and have equivalent risks of misstatement; and (ii) a single component could be 100% misstated. In the second example, we relax the assumption of equivalent size/risk. In the third example, we relax the assumption that a single component may be 100% misstated, and consider other possible error patterns. Following these examples, we then discuss audit procedures to be undertaken at selected components and conclude with possible extensions.

4.1 | Example 1: Equal-size components

In the simple situation of equal size components, two critical items of information are relevant. The first is the number of total components, and the second is the number of components that would need to be mostly or totally misstated (i.e., the “critical event”) to aggregate to an overall material misstatement. We use the second piece of information to set the minimum number of components that we must audit more closely in order to be assured, at the desired level of selection risk/assurance, that at least one example of this condition will exist

TABLE 1 Some factors influencing desired overall assurance

	Higher assurance required	Lower assurance acceptable
Centralization	Decentralized	Centralized, unless a risk of management fraud
Internal controls	Controls are only satisfactory or are not tested as a basis for reliance	Well controlled
Internal audit	Not present, or work not relevant	Work supports audit objectives
Proximity of components	Risk of shifting inventory	Little risk of shifting inventory

and be detected at a selected component, thus alerting the auditor to the potential problem and signaling a need for expanding the planned extent of audit procedures.[17]

To illustrate, we employ an example based on a large retail chain store audit. We first assume equal asset size components, a reasonable assumption, especially after treating separately any very large or risky and insignificant asset size components in the preliminary step outlined earlier. Certain retail outlets, for example, often must attain a certain size to be efficient, and may be more or less similar in size. Assume this company is not a new client. Controls over the existence and valuation objectives for this decentralized company are considered effective. Past audit results have not indicated any significant misstatement issues or audit areas. The threshold set for the existence objective for the components is 10% of the $250,000,000 total asset balance, or $25,000,000.

The process is straightforward when all components are of approximately the same size, the a priori error patterns and error rates are also assumed to be similar (the likelihood of misstatement is the same at each component), and it is possible that a component might be 100% misstated. This situation is illustrated in Table 2. To determine the minimum number of components to be audited, compute the ratio of the number of components that must contain the critical event (e.g., a 10% or greater misstatement) to add to the threshold for the test.[18] If each component is about $1,000,000 in value (i.e., 250 components adding to $250,000,000), then it would require 25 "bogus" (100% misstated) components to aggregate to a 10% overall misstatement of the assets. By the use of software, formulae, or tables, the auditor can determine the minimum number of components to be examined by finding how many components would need to be examined to reveal at least one instance of this misstatement pattern at a specified level of confidence.

Appendix B shows three example tables derived using the attribute sampling module in IDEA.[19,20] These tables illustrate selection risk assurance levels of 51%, 80%, and 90%, and various "critical event" percentages for populations of 30, 100, and 250 components.[21] The confidence levels of the tables relate to the desired selection risk from the selection of components. Separate tables can be developed for different population sizes. Interpolation can often be used to determine testing levels when the engagement circumstances create situations that fall between the table values when exact computations using the hypergeometric distribution are not available. Such tables also illustrate the interactions between population size and other sampling parameters, such as risk and the established threshold. In the current example in Table 2, the auditor would need to identify at least 15 components in order to achieve an 80% (moderately high) level of selection assurance that at least one instance of a "critical event" component will appear in a random sample. The 80% selection assurance is supported by inherent and control risk assessments and an assumption that analytical procedures will be applied to the components.

The procedure illustrated here is based on the "discovery sampling" technique of statistical sampling, which is captured mathematically in Appendix A. Its purpose is to detect a totally or mostly misstated component or find an instance perhaps indicating a pattern of misstatement, not to set a value for the population. Thus, if even *one* instance of this condition is found in the sample, audit action and rethinking of the risks is required. Even though misstatements of consequence and patterns of systemic error may be infrequently encountered in practice, any situation discovered that reflects such a possible risk or error pattern warrants closer attention to assess its significance. Based on guidance in auditing standards, the auditor should investigate the reasons for all discrepancies identified, even if corrected by the client. Discovery of misstatements may trigger an

TABLE 2 Controlling component selection assurance/risk

Number of components	Size	Value
250		$250,000,000
Selection assurance worksheet		
Description	Number	Value
Total components	250	$250,000,000
Examined—size/risk; insignificant	(0)	(0)
Remainder = Population for testing	250	$250,000,000
Overall threshold for test objective		$25,000,000
Less: amount set aside—nonsampled components[a]		(0)
Remainder = Threshold		$ 25,000,000
Threshold as a percentage of test population dollars		0.1 (10%)
Critical event components—summing to threshold		
(a) equal size assumption (e.g., 250 × .1)		25
Critical event components as a percentage of total components[b]		10.0%
Desired selection risk assurance		80%
Minimum components		
(a) by exact computation		15
(b) by tables (see Appendix B)		15

[a]Individually examined locations due to risk or size or insignificance.

[b]In units (25/250 = 10%).

[c]Using IDEA (Audimation, Version 10, 2017).

expansion of the plan to other, previously unselected components, to reduce the risk that an observed error condition or pattern is more widespread. More precise computations of risk can be made statistically or judgmentally once a pattern of error is identified.

There are at least two possible responses to the statistical interpretations for the "critical event," and they are similar to the interpretations when making an inference from a hypothesis test. One is that the critical event may be an indication that there could be an overall material misstatement. The hypothesis-testing equivalent would be correctly rejecting a null hypothesis. The other is that we have stumbled upon an unusual component that is not representative; there will not be many (if any) other such components, and thus there may not be a potential material misstatement. The "good news" (statistically) is that *not* finding the defined critical event supports that we examined enough components to be able to conclude at the desired level of selection risk. Designing this phase to detect a possible pattern of 100% misstatement would also likely detect many other possible misstatement patterns, which would need to be considered by the auditor before concluding that enough components have been identified.[22]

In summary, the implication of finding a critical event is to investigate further when any instance of the critical event is identified in a component. Even when a critical event is not identified, some misstatement may be found in the components examined. Sampling expertise may be helpful in analyzing the pattern of any errors of consequence found, and reassessing the possible corresponding risks. Again, professional judgment in conjunction with statistical theory can help identify the characteristics of components that are particularly likely to exhibit material errors.

4.2 | Example 2: Unequal size and/or risk

In our second example, we relax the assumption that all components are of equal size. In such situations, a conservative assumption can be used to identify the minimum number of components that would need to be fully or mostly misstated by assuming the largest components are the ones misstated (of course, this would be after removing large, high-risk or full-scope audit components). One procedure that efficiently achieves this purpose is to cumulatively sum the surrogate value of the components, starting with the larger ones until the materiality threshold is reached. This assumption effectively will lower the number of "critical event" components the auditor is trying to find, and raise the number of components to be examined accordingly. Table 3 illustrates the application of this approach to a population of components stratified by size.

In the Table 3 example, the asset value across the 250 components ranges in size from $3,000,000 to $500,000. This table illustrates the common pattern of a few very large components and many smaller ones. Thus, it is efficient to count through the component values when they are arrayed in descending value order. The threshold value of $25,000,000 is reached in this population within the tenth component.[23] Rounding down to the ninth component provides a conservative target for estimating the minimum number of components to be audited. Following the procedures outlined in Table 3, the threshold rate of misstatement becomes 3.6% of the population size and results in the required number of components for audit attention of 37 for the same level of selection confidence as in the first example.

A strategy to possibly reduce the overall number of components audited in this sampled population would be to assign one or more of the larger components in the population to the group of components already selected to be audited due to their size or risk, and recalculate the number of components that could add to the threshold materiality. This sometimes results in fewer overall components being audited.

4.3 | Example 3: When 100% misstatement is unrealistic

While the most conservative results are obtained by assuming that the errors are seeded in the fewest number of components (e.g., some components are 100% misstated—in essence an existence issue), in many situations the auditor will believe that this "worst case" scenario (100% misstatement) is unrealistic. Suppose the auditor judges that any misstatement above 60% of the value of a component could be easily detected by other signals, such as a lack of remitted earnings or procedures based on analytics and other factors (e.g., results of internal audit procedures). The use of such an assumption in designing tests is adapted from the guidance in AICPA audit guide *Audit Sampling* in a section entitled "Designing samples to address assertions" (e.g., AICPA, 2014, paragraph 4.59).

Table 4 illustrates applying such a judgment, using otherwise similar parameters to those used in Table 3. The minimum number of components necessary to contain one example of a revised critical condition and "trip" the established threshold (considering the unequal size condition) increases to 18 components from the Table 3 example of nine components, and the minimum number of components to audit declines from 37 to 20. However, this reduction in the number of components is crucially dependent on the support for the 60% maximum assumption. Care must be taken to not routinely "assume away" the risks, and thus reduce the components audited, as there are many examples of costly mistaken judgments made by auditors concerning the likelihood of misstatements. Such judgments should be supported by evidence and reconfirmed each audit period.

4.4 | Controlling detection risk at the selected components

In some situations, there can be a low audit risk of failing to detect a critical misstatement condition when the auditors perform a full scope or statutory audit at the component for purposes of reporting separately on the component or to meet local country regulatory requirements. In such cases, a relatively high level of achieved assurance is indicated on the accounts, balances and assertions regarding that component, and the major risk becomes the selection risk of the number of components to examine.[24] If the assurance that the auditor is expected to attain from the component procedures is very high, it may help to identify other situations where more risk may be indicated, such as:

TABLE 3 Controlling component selection risk

Number of components	Size	Value	Cumulative value
3	$3,000,000	$9,000,000	$9,000,000
5	$2,500,000	$12,500,000	$21,500,000
10	$2,000,000	$20,000,000	$41,500,000
10	$1,500,000	$15,000,000	$56,500,000
202	$908,416	$183,500,000	$240,000,000
20	$500,000	$10,000,000	$250,000,000
Total = 250 components		$250,000,000	
Selection assurance worksheet			
Description		Number	Dollars
Total components		250	$250,000,000
Examined—size/risk; insignificant		(0)	(0)
Remainder = Population for testing		250	$250,000,000
Overall threshold for test objective			$25,000,000
Less: set aside for nonsampled components			(0)
Remainder = Threshold			$25,000,000
Threshold as a percentage of testing population dollars			0.1 (10%)
"Critical event" components—summing to threshold			
(a) equal size assumption (e.g., 250 × .1)			25
(b) unequal size—based on count[a]			9
(c) smaller of (a) or (b) above			9
Critical event components as a percentage of total[b]			3.6%
Desired selection risk assurance			80%
Minimum components			
(a) by exact computation[c]			37
(b) by tables (see Appendix B)			~43

[a]Based on a stratification like the one at the beginning of the example, count the number of 100% misstated components that would be needed to aggregate to a material misstatement. In this case, the threshold is reached while within the tenth location. For conservatism, round down (e.g., nine). This is easily accomplished with a data sort on location size first, and then cumulative subtotaling each location. When some components in this population are large, it may be more efficient to separate them out before sampling.

[b]In units (9/250 = 3.6%).

[c]Using IDEA (Audimation, Version 10, 2017). Differences with the tables can be attributed to the ability to treat the sampling parameters as continuous rather than discrete values.

- unexpected deficiencies in the effectiveness of internal control procedures;
- local differences in the control environment and attention to controls monitoring.

In those situations where less than a full-scope audit is planned, and detection risk needs to be considered at the component level, the extent of assurance should be consistent with the desired overall assurance (i.e., Overall assurance = Selection assurance × Detection assurance). For example, to achieve an overall assurance of around 81% from testing and with a selection risk of 10% (selection assurance of 90%), the risk computation indicates that each component should be audited to detection risk of 10% or less (0.81/0.9 = 0.9).

Auditors using random selection to identify the components audited, and also employing random sampling to identify items for testing when a sample of them is examined, can employ classical (or monetary unit) statistical sampling formulae to project and evaluate the statistical limits (bounds) of any testing results. Judgmental samplers will not be able to state with certainty a specific achieved precision and confidence for the test, but auditors should still apply judgment in deciding if the planning assumptions are still valid and the test results are acceptable within the limits of component materiality, group performance materiality, and group materiality or indicate further investigation or evidence is necessary.

4.5 | Relating component materiality to the formulation

ISA 600 requires the auditor to determine component materiality for the components examined. Component materiality should be less than group materiality, and the auditor may consider various factors in determining how much less than group materiality is warranted under the circumstances. This measure can be determined by judgment and would, in combination with the detection assurance/risk sought via use of the formulation in this paper, suggest a level of audit procedures to meet the parameters of risk and component materiality at the selected components. When the component materiality determined by judgment and the threshold used in the formula differ, the lesser

TABLE 4 Controlling component selection risk

Number of components	Size	Value	Cumulative value
3	$3,000,000	$9,000,000	$9,000,000
5	$2,500,000	$12,500,000	$21,500,000
10	$2,000,000	$20,000,000	$41,500,000
10	$1,500,000	$15,000,000	$56,500,000
202	$908,416	$183,500,000	$240,000,000
20	$500,000	$10,000,000	$250,000,000
Total = 250 components		$250,000,000	
Selection assurance worksheet			
Description		Number	Dollars
Total components		250	$250,000,000
Examined—size/risk; insignificant		(0)	(0)
Remainder = Population for testing		250	$250,000,000
Overall threshold for test objective			$25,000,000
Less: set aside for nonsampled components			(0)
Remainder = Threshold			$25,000,000
Threshold as a percentage of testing population dollars			0.1 (10%)
Worst-case error assumption			0.6 (60%)
Critical event components—summing to threshold			
(a) equal size assumption (e.g., 250 × .1/.6)			42
(b) unequal size—based on count[a]			18
(c) smaller of (a) or (b) above			18
Critical event components as a percentage of total components[b]			7.2%
Desired selection risk assurance			80%
Minimum components			
(a) by exact computation[c]			20
(b) by tables (see Appendix B)			~22

[a]Based on a stratification like the one at the beginning of the example, count the number of 60% misstated locations that would be needed to aggregate to a material misstatement. In this case, the threshold is reached while within the 19th location. For conservatism, round down (e.g., 18). This is easy with a data sort on component surrogate value first, and then cumulative subtotaling each component value. When some components are large, it may be more efficient to separate them out before sampling.

[b]In units (18/250 = 7.5).

[c]Using IDEA (Audimation, Version 10, 2017).

value should conservatively be used in satisfying detection risk at the component level.

4.6 | Extensions of the formulations in this paper

Once the number of components to be examined has been determined, the auditor may be able to improve the probability of finding a component of audit interest by applying factor weightings to the components based on their differing risk characteristics, as discussed in the general model in Appendix A. For example, a component that has not been audited for some time might carry a higher a priori risk assessment and warrant a greater likelihood of selection. Such weightings can also be reflected in the estimates of the "worst case" situations by setting those assumptions for each component or a stratum of more similar risk components based on known factors and judgments. In such cases the cumulative number of components necessary to satisfy the risk targets would reflect both the component size and its individual worst case assumption.

When strata can be identified by differing risk levels associated with components, one approach would be to apply risk weightings to the surrogate values of the component items in the various strata, and allocate the indicated component sample size to the strata in proportion to their "risk-weighted" sum values. This method would provide a more risk-sensitive selection, compared with simple random selection. While this would introduce some bias in the selection, it would do so in favor of detecting a critical condition if it exists and thus provide a more conservative approach than when not factoring in the weightings.

When the cost of auditing at a component (e.g., travel cost and lost travel time, local orientation and any required training) is added to the formulation, the auditor may be able to additionally control risks and minimize the associated costs of auditing. Such extensions are beyond the scope of this paper, but should be formally considered in further research.

5 | CONCLUSIONS, LIMITATIONS, AND FUTURE DIRECTIONS

As businesses continue to become more complex and global, growth and mergers create opportunities for businesses to paint financial pictures that meet expectations, but not reality. As auditors respond to new standards and regulatory requirements regarding their responsibility to detect material misstatements, the profession is further challenged to find a practical way to address and document the thought process of planning such engagements. Economics demands that auditors seek a level of work that demonstrates professional standards, and is both efficient and risk sensitive. The historical and continuing number of litigations and inspection and peer review issues encourage the search for better solutions to this significant and widespread issue.

To date, there has been limited guidance available to auditors to assist in the quantitative aspects of planning the component audit. Auditors have long struggled with the issue of how to control audit costs, while performing sufficient verification to ensure that risk of material misstatement is controlled. The availability of guidance to assist professional judgment in the audit planning phase should help auditors to demonstrate the reasonableness of the work performed, and to document and communicate the factors of audit judgment that were used in planning. Setting the scope of components to be audited is especially important when demonstrating low audit risk in an environment where a large number of audit components exist, and when data and supporting evidence are not accessible in a central location. In many entities with many components, evidence supporting the component financial statements needs to be gathered at the component site. The method illustrated for controlling selection and detection risk may be helpful in this process, by providing guidance in setting minimum number of audit components and extent of procedures to be performed at the component level to achieve a desired level of audit assurance.

An advantage of the approach illustrated here is that it does not require the auditor to assume all components will be audited and thus allocate audit effort to each nontrivial component. When it is practical and possible to audit all components, existing approaches to allocating work effort to the components can be used, such as those proposed by Dutta and Graham (1998), Glover et al. (2008), or Stewart and Kinney (2013). The formulation in this paper provides guidance in determining the minimum number of components to be subject to audit procedures for a desired level of audit assurance that a "critical" misstatement or pattern of misstatement would be detected.

However, there are associated limitations. While the procedure is a decision aid based on statistical principles, the number of judgments involved indicates that the implied answer is itself a judgment, rather than a purely mathematical result. In situations where extensive experience with the client and the associated risks of misstatement exist, more precise calculations based on traditional classical statistical sampling plans[25] can be made.

Further, not all the potentially relevant inputs have been formally incorporated in the examples here. Specific future directions for this research are the inclusion of specific judgmentally weighted variables in the general model to better target risks in specific components.[26] These weights can be formally recognized in the modeling methodology. Such modeling extensions can be useful in developing a generalized optimization model for risk or cost minimization, but are beyond the scope of this paper.

Note that even if scoping decisions adequately address potential selection risk, the incorrect application of component materiality concepts could still increase the likelihood of missing a material misstatement. As such, addressing selection risk by itself is not sufficient to reduce group audit risk overall. Further, we do not explicitly describe a single method for extrapolating component sample results to the population, as this will be a function of how the sample was selected. Once the sample selection strategy is determined, the extrapolation would follow appropriate projection techniques for that strategy.

While we provide a potential way forward in resolving a long-standing conundrum, there remain interesting questions for future study on the frontiers where auditing and statistics meet. These include:

- Could auditors over time develop models for specific businesses or even whole industries that allow us to lower selection risk by identifying high-risk situations in advance?
- Can auditors effectively use ensemble methods (Seni & Elder, 2010; Zhu, 2008) to combine judgment and experience with statistical models to improve efficiency? This question is being addressed in fields as diverse as sports team management (the "Moneyball" effect) and financial trading (balancing quantitative models with expert opinion). In the component environment, an effective ensemble approach would balance the risks assigned by an auditor's judgment and experience with a weighted randomization consistent with classical statistical or monetary unit sampling.
- Can existing methods of effort allocation (e.g., Glover et al., 2008; Stewart & Kinney, 2013) be combined with the formulation proposed here to enhance the usefulness of these models?
- While much has been written on the subject of components, the judgment regarding what constitutes a component remains elusive in practice. Will further experience and guidance help refine and make auditor judgments more consistent?
- What is the nature and extent of educational efforts and practice aids that will be necessary to implement these approaches?

ENDNOTES

1 PCAOB inspection reports often identify deficiencies in sufficiency of evidence in group audits; for example, the 2014 inspection report of Ernst & Young states, "(t)he Firm … failed to assess the risks of material misstatement associated with the individual locations and use that assessment to determine the locations at which to test these controls" (PCAOB, 2015, p. 16). Further, the Accounting and Corporate Regulatory Authority (2016) of Singapore notes that audit quality issues are "often due to inadequate audit evidence retained in the group audit working papers on the extent of testing performed by the component auditors" (p. 40), and "inappropriate scoping of significant components" (p. 55).

2 Paragraph .11 of AU-C section 600 *Group Audits* (AICPA, 2012d) defines a component as: "An entity or business activity for which group or component management prepares financial information that is required by the applicable financial reporting framework to be included in the group financial statements." The diverse locations (components) in a multilocation entity could fall under this definition. Since this terminology is new and may relate to locations, the term *locations* is also used in this paper to assist readers in relating the concept.

3 Literature prior to AU-C section 600 *Group Audits* generally referred to the problem addressed in this paper as the "multilocation" problem. We use both the terms "location" and "component," realizing that in some situations locations and components will be the same and in other situations they will differ.

[4] Phar-Mor, a large multiple-component drug store chain, committed an inventory fraud in which the extent of overstatement of ending inventory grew over time, and went undetected for several years. When detected through a "tip," the ultimate loss to third parties approached a half-billion dollars. Losses to auditors Coopers & Lybrand in litigation and settlements were reported to be in excess of this amount. As stated by the auditor's counsel, "… the (Phar-Mor) executives made bloody sure that there was nothing wrong with the inventory in the stores we checked. The manipulation was in the stores we didn't go to" (Stern, 1992). The few locations where auditor inventory observation took place made the manipulation of inventory easier.

[5] The concept of specifying that component materiality be less than group materiality is derived from statistical theory for planning and evaluating independent sample results from different strata or sampling applications (see AU-C section 600 *Group Audits)*.

[6] The PCAOB provided some qualitative guidance on how to structure component audits of internal controls in auditing standard no. 2 (PCAOB, 2004), and later expanded that guidance in auditing standard no. 5 (PCAOB, 2007, paragraph 33 and appendix B). However, neither standard provides a quantitative approach to allocating work effort to components.

[7] Under the clarity standards, Paragraph .14 of *AU-C Section 200: Overall Objectives of the Independent Auditor and the Conduct of an Audit in Accordance with Generally Accepted Auditing Standards* defines detection risk as: "The risk that the procedures performed by the auditor to reduce audit risk to an acceptably low level will not detect a misstatement that exists and that could be material, either individually or when aggregated with other misstatements" (AICPA, 2015). In this paper, detection risk is focused on a specified misstatement condition defined by the auditor for the component.

[8] Performance materiality is a new term under the Clarity standards (*AU-C Section 320: Materiality in Planning and Performing an Audit*) and is an amount less than materiality and allows for the aggregation of account-level conclusions to support a low overall audit risk (AICPA, 2012b). For a further explanation of the relationships between various concepts of materiality in the planning of Group Audits, see *AU-C Section 320* and particularly *320.32c*. Tolerable misstatement is a concept used in designing audit tests that then aggregate to account level conclusions and overall. Its formal definition is: "A monetary amount set by the auditor in respect of which the auditor seeks to obtain an appropriate level of assurance that the monetary amount set by the auditor is not exceeded by the actual misstatement in the population" (AICPA, 2012c, AU-C section 530.05). Tolerable misstatement may be less than performance materiality or may be set to the same level.

[9] We note two other recent papers on audit sampling. First, Elder, Akresh, Glover, Higgs, and Liljegren (2013), a synthesis of past sampling research, mentions appendix E on multilocation auditing in the *Audit Sampling* guide, acknowledges the lack of information regarding this procedure, and suggests further research. Durney, Elder, and Glover (2014) focus on rates of misstatement in tests of the accounts and balances in general. While they conclude that sampling procedures have improved since the Sarbanes–Oxley Act (U.S. Congress, 2002), they do not specifically focus on complications specific to group audits.

[10] Allen and Loebbecke (1998) interviewed six experienced auditors and reviewed four firm audit manuals, but did not examine actual audit decisions evidenced in work papers.

[11] For example, as in the audit of a multinational enterprise like Coca-Cola, or a retailer like Home Depot.

[12] For example, in the Olympus fraud, transactions with related party entities helped hide corporate losses from view. In the absence of the auditor seeing the transactions and their implications from both the main entity and the related party entity perspective, the true impact of these transactions was difficult to detect (Dutta, Caplan, & Marcinko, 2014).

[13] In many different applications, statisticians are increasingly relying on "ensemble methods" (Seni & Elder, 2010; Zhu, 2008) to form an overall opinion based on results from several competing models with different assumptions. While such methods are beyond the scope of this paper, auditors might make similar inroads with a combination of judgment and statistical techniques, driven by the risk profile of a particular client.

[14] Risk and assurance are complementary: 5% risk implies 95% assurance. The relationship here is as described in appendix E of the audit guide *Audit Sampling* (AICPA, 2014).

[15] Note that, using statistical theory, a measure of variability of the characteristic of interest (e.g., audited value, differences from recorded value, and ratios of audited to recorded value) is used to evaluate sample results and determine statistical limits and combine limits (or plan for limits). Since such a measure is not feasible to use for purposes of selection, a recorded value of some characteristic of the components is used as a surrogate for the determination of components for audit effort. This is also the approach for the allocation of materiality illustrated in Glover et al. (2008).

[16] The general model developed in Appendix A also allows for patterns of errors to vary across components.

[17] The identification of such a condition should indeed be a rare event if the entity is represented as having an effective system of internal control and a well-functioning control environment. However, identifying such a condition or critical event could be highly important to identifying a potentially risky audit situation.

[18] Tolerable misstatement may be thought of as a materiality threshold applied at the level of the sample or test within an account.

[19] IDEA (Audimation Inc.) is a commonly used auditing and sampling tool that uses the hypergeometric distribution to compute sample sizes. See Stewart (2012) for an explanation of the use of the hypergeometric distribution using an Excel function.

[20] The use of tables to provide guidance in large populations on the extent of testing is well established. For example, the AICPA (2014) provides table values for determining sample sizes for audit sampling. However, AICPA's tables rely on the binomial distribution, assuming sampling with replacement. However, auditors sample *without* replacement, implying that the hypergeometric distribution is appropriate. The error caused by violating the assumptions of the binomial distribution is low in large populations, but increases as population size declines. The number of components for an entity would commonly be low enough so as to cause significant differences between the values in the *Audit Sampling* guide and those obtained from the hypergeometric distribution used in our method. For example, the tables in the *Audit Sampling* guide are only designed for large populations and confidence levels of either 90% or 95%. Readers can compare the relative sample sizes in Appendix B for the three population sizes and for various confidence levels to note the limitations of these tables for purposes of this formulation. Even with larger populations, commonly used software such as ACL (Audit Command Language @ acl.com) which uses the binomial distribution, using the same parameters, computes the number of components in Example 1 to be 17, rather than the 15 as noted in the example which uses the hypergeometric distribution.

[21] While information on the number of components in actual group audits is rare, Downey and Bedard's (2017) survey evidence shows a maximum number of 54 components in a nonrepresentative sample of recent Big 4 global group audits. Thus, we illustrate situations with both fewer and more components than found in their sample.

[22] The use of the "worst case" assumption here reduces the number of "critical events" in the population and increases the number of components audited in order to detect one example. Weighting selection probabilities for qualitative issues would be an extension of this formulation, as discussed in Appendix A.

[23] After the tenth component or location, the cumulative dollars equal $25,500,000.

[24] The criteria often associated with a full-scope audit would ordinarily be more conservative (lower) than the scope of any allocated materiality from the group audit perspective.

[25] Stratified classical statistical sampling using normal or another distribution theory.

[26] For example, the time since the component was last closely examined, recent changes in key employees and management, variances from

expected accounting values, and evidence of recent past operating or accounting problems, among other factors, can be used to weight the various components in order to improve the detection ability of the approach. Differing costs of visiting or auditing various components and any differing cost of auditing at a selected component is another factor that could be more formally incorporated in an expanded model.

ORCID

Jean C. Bedard http://orcid.org/0000-0001-9691-8040

REFERENCES

Accounting and Corporate Regulatory Authority. (2016). Practice monitoring programme public report. Retrieved from: https://www.acra.gov.sg/Publications/Reports/Practice_Monitoring_Programme_Public_Reports.

Allen, R., & Loebbecke, J. (1998). A risk model for multilocation audits. *Journal of Applied Business Research*, *14*(4), 1–13.

American Institute of Certified Public Accountants. (2012a). *AU-C section 240: Consideration of fraud in a financial statement audit*. New York, NY: AICPA.

American Institute of Certified Public Accountants. (2012b). *AU-C section 320: Materiality in planning and performing an audit*. New York, NY: AICPA.

American Institute of Certified Public Accountants. (2012c). *AU-C section 530: Audit sampling*. New York, NY: AICPA.

American Institute of Certified Public Accountants. (2012d). *AU-C section 600: Special considerations—audits of group financial statement (including the work of component auditors)*. New York, NY: AICPA.

American Institute of Certified Public Accountants. (2013). *Risk alert. Understanding the responsibilities of auditors for audits of group financial statements*. New York, NY: AICPA.

American Institute of Certified Public Accountants. (2014). *Audit guide: Audit sampling*. New York, NY: AICPA. (Note: Appendix E of this guide was originally published in the 2008 version of the guide and has been carried forward to subsequent editions.)

American Institute of Certified Public Accountants. (2015). *AU-C section 200: Overall objectives of the independent auditor and the conduct of an audit in accordance with generally accepted auditing standards*. New York, NY: AICPA.

American Institute of Certified Public Accountants. (2016). *Audit guide: Assessing and responding to audit risk in a financial statement audit*. New York, NY: AICPA.

Asare, S. K., Fitzgerald, B. C., Graham, L. E., Joe, J. R., Negangard, E. M., & Wolfe, C. J. (2013). Auditors' internal control over financial reporting decisions: Analysis, synthesis, and research directions. *Auditing: A Journal of Practice & Theory*, *32*(Supplement 1), 131–166.

Christensen, B. E., Elder, R. J., & Glover, S. M. (2015). Behind the numbers: Insights into large audit firm sampling policies. *Accounting Horizons*, *29*(1), 61–81.

Cochran, W. G. (1977). *Sampling technique* (3rd ed.). New York, NY: John Wiley & Sons, Inc.

Downey, D. H., & Bedard, J. C. (2017). Coordination and communication challenges in global group audits. Working paper, Villanova University and Bentley University.

Durney, M., Elder, R. J., & Glover, S. M. (2014). Field data on accounting error rates and audit sampling. *Auditing: A Journal of Practice & Theory*, *33*(2), 79–110.

Dutta, S. K. (2013). *Statistical techniques for forensic accounting: Understanding the theory and application of data analysis*. Upper Saddle River, NJ: FT Press.

Dutta, S. K., Caplan, D. H., & Marcinko, D. J. (2014). Blurred vision, perilous future: Management fraud at Olympus. *Issues in Accounting Education*, *29*, 459–481.

Dutta, S. K., & Graham, L. E. (1998). Considering multiple materialities for account combinations in audit planning and evaluation: A cost efficient approach. *Journal of Accounting, Auditing and Finance*, *13*, 151–171.

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

Elliott, R. K., & Rogers, J. R. (1972). Relating statistical sampling to audit objectives. *Journal of Accountancy*, *134*, 46–55.

Glover, S. M., Prawitt, D. F., Liljegren, J. T., & Messier, W. F. Jr. (2008). Component materiality for group audits. *Journal of Accountancy*, *206*, 42–46.

International Audit and Attest Standards Board. (2009). *International standard of auditing no. 600. special considerations – audits of group financial statements (including the work of component auditors)*. New York, NY: IFAC.

Kim, H. S., Neter, J., & Godfrey, J. T. (1987). Behavior of statistical estimators in multilocation audit sampling. *Auditing: A Journal of Practice and Theory*, *6*, 40–58.

Lemon, W. M., Tatum, K. W., & Turley, W. S. (2000). *Developments in the audit methodologies of large accounting firms*. Caxton Hill, UK: Stephen Austin & Sons Ltd.

Panel on Audit Effectiveness. (2000). *Report and recommendations*. Stamford, CT: Public Oversight Board.

Public Company Accounting Oversight Board. (2004). *Auditing standard no. 2: An audit of internal control over financial reporting performed in conjunction with an audit of financial statements*. Washington, DC: PCAOB.

Public Company Accounting Oversight Board. (2007). *Auditing standard no. 5: An audit of internal control over financial reporting that is integrated with an audit of financial statement*. Washington, DC: PCAOB.

Public Company Accounting Oversight Board. (2010). *Auditing standard no. 8: Audit risk*. Washington, DC: PCAOB.

Public Company Accounting Oversight Board. (2015). *Report on 2014 Inspection of Ernst & Young LLP*. PCAOB, Washington, D.C.

Public Company Accounting Oversight Board. (2016). Proposed amendments relating to the supervision of audits involving other auditors and proposed auditing standard—dividing responsibility for the audit with another accounting firm. Concept release no. 2016–002. Washington, DC: PCAOB.

Roberts, D. M. (1978). *Statistical auditing*. New York, NY: AICPA.

Sanders, D. (2014, August 12). Qualitative considerations for allocating materiality to components in a group audit. *Journal of Accountancy*. Retrieved from https://www.journalofaccountancy.com/news/2014/aug/group-audits-201410712.html (accessed September 20, 2017).

Securities and Exchange Commission. (2002). *Rite Aid's former management team charged with securities fraud by SEC*. Litigation release no. 17577. Retrieved from www.sec.gov/litigation/litreleases/lr17577.htm (accessed December 2016).

Seni, G., & Elder, J. (2010). *Ensemble methods in data mining: improving accuracy through combining predictions*. Morgan & Claypool Publishers.

Stern, G. (1992, August 18). Phar-Mor seeks Chapter 11 protection from creditors. *The Wall Street Journal*.

Stewart, T. (2012). *Technical notes on the AICPA audit guide Audit Sampling*. Version 1.0. AICPA. Retrieved from https://www.aicpa.org/Publications/AccountingAuditing/KeyTopics/DownloadableDocuments/Sampling_Guide_Technical_Notes.pdf (accessed September 20, 2017).

Stewart, T., & Kinney, W. (2013). Group audits, group-level controls, and component materiality: How much auditing is enough? *The Accounting Review*, *88*, 707–737.

Sunderland, D., & Trompeter, G. M. (2017). Multinational group audits: Problems faced in practice and opportunities for research. *Auditing: A Journal of Practice & Theory*, *36*(3), 159–183.

U.S. Congress. (2002). *The Public Company Accounting Reform and Investor Protection Act of 2002 (The Sarbanes–Oxley Act)*. Pub.L.107–204, 116 Stat. 745 (July 30). Washington, DC: US Government Printing Office.

Westervelt, M. A. (2014, February 18). The scoop on group audits: You may have them, even though you think you don't. *Journal of Accountancy*. Retrieved from http://competency.aicpa.org/media_resources/208476-the-scoop-on-group-audits-you-may-have-them-even/detail (accessed September 20, 2017).

Zhu, M. (2008). Kernels and ensembles. *The American Statistician*, *62*, 97–109.

Lynford Graham, Visiting Professor at Bentley University, is a certified public accountant with more than 30 years of public accounting experience in audit practice, national policy development groups of major firms, and setting auditing standards. His research interests span a number of topics related to improving auditor judgment and audit quality. He is presently an author, educator, and consultant, with professional assignments focusing on multidimensional problem solving requiring "leading edge" thinking and "hands-on" management.

Jean Bedard, Timothy B. Harbert Professor of Accountancy at Bentley University, has published widely in scholarly and professional journals on topics related to audit quality. Professor Bedard has served as Vice President-Publications of the American Accounting Association, and has also been President, Secretary, and Historian of the American Accounting Association's Auditing Section. Her work was recognized by the Auditing Section with the Outstanding Auditing Educator Award (2012) and the Distinguished Service Award (2017).

Saurav Dutta, Head of School at Curtin University, Australia, received his Bachelor of Technology in Aerospace Engineering from the Indian Institute of Technology, Bombay and his PhD in Accounting from the University of Kansas. He has served as an Academic Fellow at the Securities and Exchange Commission, Washington, DC. He has provided statistical expertise in many high-profile corporate litigation cases.

How to cite this article: Graham L, Bedard JC, Dutta S. Managing group audit risk in a multicomponent audit setting. *Int J Audit*. 2017;1–15. https://doi.org/10.1111/ijau.12103

APPENDIX A

A generalized mathematical model of a multiple-component problem

In this appendix we develop a generalized model of a multiple component (multilocation) audit problem drawing from a two-stage sampling technique. Two-stage sampling is recommended for situations in which "tests can be performed on a small amount, … likely to be drawn as a sub-sample from a larger amount that is itself a sample" (Cochran, 1977, p. 274). Suppose multiple components are denoted as L_1, L_2, …, L_n. Each component has a surrogate amount denoted as λ_1, λ_2, …, λ_n. The critical event (error) rate for each component is denoted by ρ_1, ρ_2, …, ρ_n, and the prior probability distribution over these rates is denoted by a probability function π. That is, $\pi(\rho_1)$ denotes the probability distribution of the critical event at component 1. In general, the amount of value in each component could be different and the probability distribution of the critical event rates could also vary across the components.

The amount of misstatement (error) expressed in dollar amount at a component ε_i is the total value times the critical event rate. A priori, the expected misstatement at any component $E(\varepsilon_i)$ is $E[\lambda_i * \pi(\rho_i)]$. The expected total misstatement in the aggregate components is

$$E(\varepsilon_t) = \sum \lambda_i * E\pi(\rho_i)$$

The estimate of the prior probability distribution of the critical event rate $\pi(\rho_i)$ is based on auditor's judgment and could vary across components. Some of the factors that influence the priors are discussed in Allen and Loebbecke (1998); additional factors are listed in other articles that identify qualitative and quantitative considerations. Values can be sorted in descending order based on their expected misstatement, (similar to Table 4), as illustrated in Table A.1. Components to audit can then be determined based on rules discussed in the paper.

TABLE A.1

Component	Size	Misstatement rate distribution	Expected misstatement	Expected proportion
L_1	λ_1	$\pi(\rho_1)$	$\lambda_1 E\pi(\rho_1)$	$\frac{\lambda_1 E\pi(\rho_1)}{\sum \lambda_i \times E\pi(\rho_i)}$
L_2	λ_2	$\pi(\rho_2)$	$\lambda_2 E\pi(\rho_2)$	$\frac{\lambda_2 E\pi(\rho_2)}{\sum \lambda_i \times E\pi(\rho_i)}$
L_i	λ_i	$\pi(\rho_i)$	$\lambda_i E\pi(\rho_i)$	$\frac{\lambda_i E\pi(\rho_i)}{\sum \lambda_i \times E\pi(\rho_i)}$
L_n	λ_n	$\pi(\rho_n)$	$\lambda_n E\pi(\rho_n)$	$\frac{\lambda_n E\pi(\rho_n)}{\sum \lambda_i \times E\pi(\rho_i)}$
Total	$\sum \lambda_i$		$\sum \lambda_i \times E\pi(\rho_i)$	

Example 2 (Table 3) in the text assumes that the prior probability distributions are equal across components; hence, the expected misstatement at each component is proportional to size. Though the strategy of picking components is discussed in terms of size, the concept is applicable to picking components in terms of expected misstatement, should the prior probability distribution differ across components. Example 1 (Table 2), in the text, is a further simplification of the aforementioned model in that the values λ_i are equal across components.

Extension to discovery sampling

In discovery sampling, presence of a single critical event in the sample would call for rejection of the premise that the population value is within an acceptable threshold. Thus, discovery of even *one* defined critical event would require conduct of additional audit procedures. However, if none of the items in the sample reveals a critical event, it provides assurance that the "critical event" (error) rate is below the minimum threshold rate. In the aforementioned model, $\pi(\rho_i)$ is the prior probability distribution of population critical events. For any ρ_i, the probability of finding zero events in a sample size of n is $\{1 - \rho_i\}^n$. When no critical event is found in the sample, the posterior probability of the event rate ρ_i, denoted as $\pi^*(\rho_k)$, is

$$\pi^*(\rho_k) = \frac{(1-\rho_k)^n.\pi(\rho_k)}{\sum_i (1-\rho_i)^n.\pi(\rho_i)}$$

If the threshold critical event rate is τ, the corresponding audit risk of discovery sampling is

$$\text{Audit risk} = 1 - \int_0^\tau \frac{(1-\rho)^n.\pi(\rho)}{\sum_i (1-\rho_i)^n.\pi(\rho_i)} d\rho$$

That is, the audit assurance of discovery sampling on not finding any critical events is the cumulative posterior distribution of error rate from zero to the threshold, or the maximum acceptable critical event rate. Note that the mathematical formulas developed in this appendix are generic, in that they make no assumptions about the underlying probability distribution of the error rates.

APPENDIX B

Tables for determining the number of components (selection assurance)

TABLE B.1 30 components

Critical percentage of components	Desired selection assurance		
	51%	80%	90%
3%	16	24	27
5%	9	17	21
10%	5	10	13
15%	4	8	11
20%	3	6	8
30%	2	4	6

TABLE B.2 100 components

Critical percentage of components	Desired selection assurance		
	51%	80%	90%
3%	17	33	44
5%	11	23	32
10%	6	13	18
15%	5	9	13
20%	3	7	10
30%	2	5	7

TABLE B.3 250 components

Critical percentage of components	Desired selection assurance		
	51%	80%	90%
3%	22	45	62
5%	14	29	40
10%	7	15	21
15%	5	10	14
20%	4	7	10
30%	2	5	7

Notes to the tables

The illustrative Tables B.1, B.2, and B.3 show the impact of the sampling parameters of risk, tolerable percentage (threshold critical percentage of components), and population size on the minimum number of components to be audited. The table values were derived using the attribute sampling module in IDEA. Exact computations in any specific situation may also be able to use the hypergeometric function in Excel—see Stewart (2012). Approximations using the binomial distribution or other functions may provide acceptable results in some circumstances—see Roberts (1978). In populations of limited size, such as those illustrated in this paper and likely to be encountered in many component sampling issues, the population size will often affect the minimum sample size. The most accurate calculations can be obtained with proper application of the hypergeometric probability distribution. Use of other distributions that are approximations of the hypergeometric or assume large populations (e.g., binomial, Poisson, or normal) may result in different (greater) minimum sample sizes.