Issues and Opportunities in Archival Audit Research

Theodore J. Mock* and Jerry L. Turner
1University of Southern California, Los Angeles, US
2The University of Memphis, US

This paper encourages audit researchers to consider archival research methodologies and assists in designing an effective research project. Archival research offers many benefits and presents many challenges not existing in other forms of research as actual audit workpapers provide a rich depiction of the actual audit process. Differences in audit methodologies and workpapers, however, create issues such as the measurement of variables and definitional inconsistency. Other issues with archival studies include relatively less control over the research process, confidentiality of data, and difficulty in obtaining audit firm cooperation. Suggestions for resolving these issues are offered and analysis of three successful archival studies is provided to demonstrate how archival research designs were adapted to the needs of the project.

Key words: Archival data, archival research, audit research, audit workpapers, research design.

INTRODUCTION

A significant knowledge gap exists in the audit literature as to various features of audit planning and decision-making as it occurs in actual practice. Archival data obtained from workpapers allow us to investigate actual audit decision-making rather than simulated judgments in an experimental setting. They also provide insights on important practice issues such as error rates in audited financial statements and identification of which audit procedures are effective in identifying risks, errors and other irregularities. Although numerous challenges exist in conducting this type of research, archival research has several significant advantages over alternative research approaches. First, the data reflect actual field results of audits and of audit decisions and judgments. Second, large repositories of potentially useful data are available in client and work-paper files. As such, the results from archival studies potentially are generalizable to real-world settings and may be more likely to influence practice and standard setters than research conducted under simulated settings. Unfortunately, although archival data allow us to address many important research questions, the potential for useful research far surpasses what actually has been accomplished.

Empirical audit research can be categorized several ways. Some empirical studies obtain data within artificial environments such as experimental studies — both laboratory and experimental-markets — and attitudinal surveys. Other studies, designated archival research, examine data resulting from actual audits.
Within this broad category are studies examining real-world data residing in public sources such as Edgar, Compustat, and U.S. Securities and Exchange Commission (SEC) documents and reports, and studies examining real-world data obtained from private sources such as audit workpapers, or from surveys related to actual audit findings. The purposes of archival audit research encompass both the general academic objectives of research — primarily the creation of general, basic knowledge — and applied research objectives such as the creation of information to aid audit judgment, practice and regulation.

Archival data residing in stock market, governmental or industry databases allow investigation of a number of audit-market research questions, such as audit quality differences between large audit firms (e.g. the Big 5) and other firms. Although this line of published research is extensive, the focus of this article is on archival data taken from audit working papers. Thus, rather than considering macro, market-level research issues, micro research issues at the individual audit and auditor levels are examined.

Archival data obtained from audit workpapers allow us to investigate actual audit decision-making rather than judgments in an experimental setting. They also facilitate research into characteristics of audits such as error rates in audited financial statements and identification of which audit procedures are effective in identifying risks, errors and irregularities.

Two important aspects of extant audit research help motivate this study. First, there exists a significant gap in reviews of audit decision-making research. For example, both Bell and Wright (1995) and Trotman (1996) place little emphasis on the potential contributions of archival research in understanding audit decision-making. Second, we believe there is a significant opportunity cost of not having real-world archival research results available to practitioners and regulators. For example, significant audit failures such as that alleged in the Lincoln Savings and Loan case (Erickson et al. 2000) possibly may have been averted if relevant field research data had been available concerning the usefulness of industry and environmental data in conducting analytical reviews. Significant opportunity costs also exist with respect to regulation, as regulators often must make decisions without scientific study of issues such as the impact of auditor rotation on audits.

Although significant challenges exist in conducting this type of research, archival research offers several significant advantages. First, as noted, the data reflect actual field results of audits and of audit decision-making. Second, large repositories of potentially useful data are available in client and workpaper files. As such, the results from archival studies potentially are generalizable to real-world settings and may be more likely to influence practice and standard setters than research conducted in laboratory settings. Unfortunately, although archival data allow us to address many important research questions, the potential for useful research far surpasses what actually has been accomplished. If some constraints to the availability of archival data can be removed or mitigated, this area of research has much to offer the profession.

The potential of archival research is not limited to practice issues as fundamental academic research questions also may be addressed. For example, research results from archival studies allow us to confirm or, in some cases, challenge or disconfirm theoretical models and findings from experimental and attitudinal research.

This paper proceeds as follows. In the following two sections, general research design and other methodological issues related to conducting an archival study are discussed. Of particular importance are issues related to obtaining cooperation from audit firms. These sections are followed by a discussion of three sample archival studies. These studies portray both the difficulties and the opportunities in conducting archival research. The final section of the paper focuses on recommendations that may help mitigate practitioners’ concerns in providing the data so essential in furthering our knowledge of auditing and assurance.

ARCHIVAL STUDY PROCEDURES

Issues about research methods for archival studies parallel those of other empirical studies in many ways, but they also exhibit some unique features. In this section we briefly discuss typical archival study design guidelines. In the next two sections, we focus on some unique design issues, in particular on issues of
obtaining audit firm participation in archival research. We assume that the reader is familiar with general research design issues such as those discussed in detail in Trotman (1996), Green and Tull (1978) and Kerlinger (1973). For a discussion of these issues, see also Trotman’s article in this publication.

One aspect of archival research is that it utilizes data arising in natural settings or what may be considered natural experiments in contrast to laboratory experiments. In terms of design, these natural experiments may result in quasi-experimental designs or, although rare, even in true experimental designs where a primary advantage is internal validity. Mock and Willingham (1983) present an example of an archival study that has most of the features of a true experimental study. Most archival studies, however, may be considered quasi-experiments as some elements of true experimental control are missing. All archival studies encompass a significant degree of external validity whether they are true or quasi-experiments. See Trotman (1996) for a discussion of the main issues and tradeoffs related to external and internal validity.

For research questions that archival studies can address, data from actual audit engagements is sought. Rather than the researcher obtaining the data directly from the working papers, however, a data gathering instrument usually is developed for auditors to abstract the desired information. If possible, the instrument is pilot tested thoroughly to ensure that the needed information is appropriately abstracted.

Having determined the research questions to be addressed, the general steps necessary to conduct an archival audit study include:

1. Identify an audit firm or firms willing to participate and liaisons with whom to coordinate.
2. Gain an understanding of the audit approach of the firm(s) and obtain and review documentation used.
3. Identify audit and client characteristics to be collected and the number of clients to be sampled.
4a. For quasi-experimental studies:
   - Design appropriate survey instruments to gather data including scaling methods for aspects not documented in the archives (workpapers).
4b. For true-experimental studies:
   - Design experimental interventions (treatments),
   - Design and collect pre-test data,
   - Select a method for random assignment,
   - Design appropriate survey instruments to gather post-treatment data including scaling methods for aspects not documented in the archives.
5. Pilot-test the data collection instruments.
6. Make necessary design changes to the instruments, distribute the instruments and monitor responses.

The instrument usually asks auditors to gather data directly from prior or current working papers. To minimize the concern that the data are generated ex post rather than being actually extracted from the working papers, the instrument may ask for specific workpaper page references where the data are found. An alternative is the audit firm can instruct staff that the instrument is a new workpaper with instructions as to which data should be recorded.

As indicated, pilot testing of the instrument should be conducted with knowledgeable practitioners. Finally, demographic data, such as staff level and years of experience regarding the audit team should be requested at the end of the instrument.

GENERAL DESIGN ISSUES

Archival studies in auditing exhibit a number of design challenges that need to be considered. These are discussed below.

Sample selection

Identifying audit clients appropriate to the research objectives can be difficult. Each large audit firm has thousands of audit clients of different sizes, from different industries, with different forms of ownership, and with different profit incentives. In addition, clients differ in the complexity of accounting issues and in the sophistication of information systems. As a rule, audit firms will not allow independent researchers access to a complete listing of clients; thus, researchers must provide guidelines as specific as possible regarding the criteria to be used by the audit firm in selecting sample companies. Among the criteria used are industry association (e.g. Bedard 1989; Mock and Wright 1993; Mock and Turner 2001), and
risk (e.g., Mock and Turner 2001). Additional criteria also may be specified, including new or continuing audit client, client size, or the existence of factors such as audit report modification.

Even though researchers may specify selection criteria, specific clients to examine generally are selected by partners in the audit firms. In some instances, clients are selected from a master list in the executive office of the audit firm. For example, Mock and Wright (1993) used a random sample from such a master list. In other instances, individual partners are selected and asked to choose appropriate clients from their individual audit client portfolios. Mock and Turner (2001) used both approaches in identifying sample clients. In some cases (e.g., Johnstone and Bedard 2001a, b) pre-selected samples must be used. No matter how the sample is selected, however, it must be recognized that there is potential for a biased sample population.

**Failure to understand or to attend to research task**

Because archival studies are not conducted in a controlled laboratory environment, the researcher cannot be certain that individuals extracting the desired information, clearly understand what the researcher is seeking. Instructions may be confusing or contain terminology unfamiliar to all participants. If information is to be extracted from specific forms or workpapers, form numbers and workpaper references may not be consistent over time, or formats may change. To identify and correct confusing or contradictory instructions, it is important that controlled pilot tests of the instrument be done and that participants in the pilot test be asked specifically about where confusion may have occurred. If feasible and necessary, training sessions on how to complete the instrument could be held with participating audit staff members.\(^6\)

A second issue in archival research is determining if participants provide sufficient attention to the task. Whether partners, managers, or lower-level staff members complete a task, the researcher can never be sure to what extent a research task was taken seriously and if the information provided represents valid data. Additionally, individual participants may be asked to answer a wide variety of queries. For some of these, he or she may have little basis for providing a response. As noted, pilot tests may help address these issues. Also, studies that involve practicing auditors in the research team (e.g., Quadackers et al. 1996 and Eimers 2001) facilitate post-data collection contacts with the participants to address data that are unclear or questionable.

**Response rate**

As noted elsewhere, an individual partner generally is under no obligation to respond to archival research requests, even when the executive office has asked for cooperation. As shown in the following table, response rates vary widely, but rates between 40 percent and 50 percent are common. Note that the three studies with the highest response rates each included practicing auditors in the research team.

Because of confidentiality issues, identifying non-respondents for subsequent follow-up can be difficult. One method of tracking used by Mock and Turner (2001) was to have survey instruments pre-numbered by the executive office of each firm and then to have the researchers report responses by that number to each executive office. To preserve confidentiality, the researchers were not given the cross-reference information between instrument number and client name. Each executive office then sent reminder notices to non-respondents.

**Definitional issues**

An issue often discovered during pilot testing is the need for definitional specificity. Audit firms may use firm-specific definitions for various aspects of an audit and these definitions may not be consistent with definitions used elsewhere. For example, Mock and Turner (2001) found that each firm participating in their study used slightly different criteria for identifying audit clients exhibiting a specific type of risk. Achieving definitional consistency across three CPA firms was important as risk was a major variable in the study.

Other terms such as materiality may require clarification as well. For example, firms may designate a planning materiality that then may be allocated between accounts. On completion of the audit, financial statements are reviewed on the basis of a materiality for the financial
Issues and Opportunities in Archival Audit Research

statements taken as a whole. The researcher must identify the terminology for materiality at different stages of an audit and make sure the appropriate definition is provided.

Measurement issues

Many factors, including the wide range of audit procedures available to auditors, create issues related to the selection of variables to examine or to the measurement of those variables. There is the problem, for instance, of how the constructs ‘extent’ and ‘nature’ of tests are measured. Is extent sample size, hours, or some other metric? Researchers sometimes must convert verbal measures into numerical values useful for statistical analysis. An audit firm, for example, may use a scale of High, Medium or Low to indicate an estimate of control risk. Using a continuous numerical scale is problematic in that there is little supportable basis for determining multiples of ‘riskiness.’ As a result, the researcher may choose to use only categorical variables to indicate the degree of estimated control risk. In other cases, auditors may be asked to indicate measures on a scale provided by the researcher. Such scales typically are 5- or 7-point Likert scales or various forms of continuous scales and may be designed more for the convenience of coding than for reflecting measures familiar to auditors. Asking auditors to respond in a manner with which they are not familiar may introduce questionable data.10

Other measurement issues often are introduced by the selection of variables. Based on professional judgment about testing a management assertion, an auditor may reduce or eliminate one audit procedure and add or increase others deemed to be more cost effective. The auditor may reduce sample size, for example, but increase testing in other areas. If sample size is a dependent variable, invalid conclusions can be drawn if those conclusions are based solely on changes in sample size. Accordingly, researchers may need to inquire about related audit program changes and budget reallocations when identifying variables of interest.

Lack of consistency across firms

Generally accepted auditing standards and their associated interpretations generally allow flexibility in how to comply with those standards. SAS No. 82 (AICPA 1997), for example, requires auditors to perform an evaluation of the risk of fraud and to document the results of the evaluation and the impact on the audit program. Mock and Turner (2001) found that each of the three participating CPA firms responded differently to this requirement. One firm designed its fraud risk documentation using the categories of risk factors identified in SAS No. 82. A second firm used both SAS No. 82 and COSO (COSO 1992) categories as a basis for evaluation, and the third firm used both SAS No. 55 (AICPA 1988) and SAS No. 82 as a basis. The impact of this variation on the research was to require the researchers to interpret each format and to place responses into common categories.

The second variation found by Mock and Turner (2001) was the form of the documentation used by CPA firms for the evaluation of fraud risk. One firm used a checklist containing the specific risk factors listed in SAS No. 82, but extended this with a substantial number of addi-

| Table 1: Response rates for selected archival studies (Source: Eimers, 2001) |
|-----------------------------|-----------------|-----------------|-----------------|
| **Study**                  | **Sample size** | **Responses**   | **Response Rate** |
| Hylas and Ashton (1982)    | 200             | 152             | 76.0%           |
| Kreutzfeldt and Wallace (1986) | 260            | 260             | 100.0%*         |
| Wright and Ashton (1989)   | 630             | 186             | 29.5%           |
| Mock and Wright (1993)     | 345*            | 159             | 46.0%           |
| Waller (1993)              | 385             | 215             | 55.8%           |
| Maletta and Wright (1996)  | 406             | 171             | 42.1%           |
| Mock and Wright (1999)     | 160             | 74              | 46.3%           |
| Eimers (2001)              | 184             | 147             | 79.9%           |

Copyright © 2001 Management Audit Ltd.  Int. J. Audit. 5: 237-249 (2001)
tional factors identified by the firm. A second firm identified only broad categories of risk factors and required the audit team to respond to open-ended questions. The third firm modified a previously existing form used to identify control strengths and weaknesses. Again, this variation required the researchers to identify a method of coding that resulted in consistent responses.

**External validity**

While many issues related to archival research involve internal validity, external validity is an issue as well. In the U.S., for example, usually only large firms agree to participate, resulting in a biased population to be studied. Further exacerbating the problem is the substantial degree of variation in audit approaches between the large firms, and the fact that in the U.S. over the past twenty years, only one of the firms (KPMG) has shown consistent willingness to participate in archival research. As a result, the external validity of most archival studies is restricted. To their credit, researchers point out that results should not be extended beyond the specific CPA firms studied, but this sometimes is overlooked when others cite results.

**ISSUES RELATED TO OBTAINING CPA FIRM PARTICIPATION**

**Confidentiality**

During the course of an audit, auditors must have access to information about personnel matters, tax matters, corporate strategy and many other types of data necessary to develop an opinion about the financial statements. Much information examined by the auditor and documented in the workpapers is considered proprietary by the audit client and client personnel trust the auditor to keep it confidential. If auditors were to reveal such information to third parties, managers would be hesitant to allow auditors to examine and document sensitive information and the integrity of the audit process would suffer. Because U.S. auditors never have been granted a privileged legal status similar to that of lawyers and doctors, the profession itself has included in a code of professional ethics, Rule 301: Confidential Client Information (AICPA 1996). This Rule of Conduct states that ‘A member in public practice shall not disclose any confidential information without the specific consent of the client.’

Compliance with Rule 301 can create problems both for the auditor and for the researcher. Because of concern over possible ethics violations and other concerns described below, US auditors are reluctant to allow academic researchers direct access to audit workpapers. They also are reluctant to ask for permission from each client that might be selected in the sample. Researchers, however, might offer several alternatives that may not violate Rule 301.

First, each member of the academic research team can sign a confidentiality agreement with the audit firm. A properly worded agreement can bring researchers under the umbrella of Rule 301 and create legal exposure for researchers if information obtained from workpapers is released inappropriately.

Researchers also can try to design data-gathering instruments and methods in ways that specific clients can’t be identified from the data provided. For example, data on each audit could be separated between ‘public’ data identifying the audit client, and ‘sensitive’ data not traceable to the audit client. The data sets would be linked via a cryptographically protected key. After a predetermined period for research analysis, participating firms would be able to revoke access to the sensitive data through destruction of the key.

For projects involving public companies, confidentiality can be quite difficult to maintain since databases such as Compustat allow searches on many different financial statement components. If even one component such as total assets is known, the identity of the client can be determined in most cases. If two or more components are known, almost any public client can be uniquely identified. This problem remains even if financial information is obtained in common-size form, such as financial ratios. However, obtaining data in bands or within ranges could reduce this possibility substantially. Also, reporting only aggregate data could help to overcome this problem.

**Legal Issues**

In addition to Rule 301, audit firms profess concern about information they share being
used against them in another forum. As one audit partner said to the authors, ‘If sharing some information could cost you $300 million in a judgment or settlement, you would be loathe to share.’ Because neither the auditors nor the researchers create privileged information, the attorneys for the plaintiff in a lawsuit potentially could subpoena information from researchers that might be more difficult to obtain from the auditors.\textsuperscript{11}

**Economic Issues**

Providing data for archival research can be quite expensive for CPA firms. Non-current files for one or more years must be identified and retrieved from physical storage. Depending on the nature of the research project, requested information may appear in many different workpaper locations. Retrieval of information often requires substantial time by partners, managers and staff — time which otherwise might be chargeable as billable hours. Using even the lowest billable rates per hour, a project requesting information from a reasonable number of clients can have an opportunity cost exceeding $100,000 U.S. Understandably, CPA firms can be reluctant to invest such substantial sums in projects which may have no clear direct bearing on their audit practice or for which potential benefits may occur only in the future.

Economics also affects when, and if, archival research projects can be accomplished. Only a few years ago, audit firms tended to have time periods where staff were 100 percent billable and other times when staff were underutilized. For external research projects willing to schedule data-gathering during slow periods, CPA firm opportunity costs were much lower, or even zero. In today’s market, however, CPA firms are able to keep staff billable virtually year-round. Accordingly, every archival research project imposes some level of opportunity cost and perceived payoffs from academic research must be more substantial.

Technological advances may be helpful in reducing the economic impact of archival research. As audit firms become more standardized with electronic workpapers, it may be possible for researchers to provide software designed to search workpaper files and automatically extract data of interest. Such software could be run ex post or could be designed to operate in the manner of a continuous audit.

**Organizational Issues**

The partnership form of organizational structure typical of U.S. audit firms and those in many other countries also results in difficulties for archival researchers. Typically, academic researchers make proposals to, and receive agreements of participation from, partners assigned to a firm’s executive office. However, executive-office partners only request cooperation from the various partners in the operating offices — each partner asked for data generally is free to decide whether or not to respond. Also, because the executive office usually does not provide a project budget to the line partners, each participating partner must be willing to absorb whatever opportunity or actual costs might be incurred. Even with close monitoring and reminders by the executive office, a response rate of 50 percent or less is typical (see Table 1).

**EXAMPLES OF ARCHIVAL RESEARCH**

Archival studies that have been conducted to date have addressed a variety of issues and provided important insights into the audit profession, as shown in Table 2. Even though these projects are published in reputable journals and are widely cited, each faced many of the issues described previously. To illustrate this, the following summaries of three archival projects highlight some of the difficulties encountered and how the researchers adapted methodologies to obtain useful data.

**Bedard (1989)**

Bedard (1989) used an archival approach to examine how and why audit plans for substantive tests are revised. A total of 54 clients of northeastern U.S. offices of three Big 8 firms participated, with survey forms completed by the audit senior in charge of the most recent audit. Restrictions as to industry and other criteria were specified, but selection of clients meeting the criteria was left to CPA firm personnel.

Most of the survey was designed to present the seniors with open-ended questions regarding audit activity in three specified
accounts for two successive audits. The open-ended format was used in an effort to avoid any influence on responses that might be induced by the presence of a list of possible audit procedures. Questions not open-ended asked the seniors to record general assessments of inherent risk for the client, as well as of the strength in the overall controls environment and in the controls for each of three accounts: accounts receivable, inventory, and accounts payable. Five-point scales were used for each assessment: ‘Very Strong’ to ‘Very Weak’ for internal control, and ‘Very High’ to ‘Very Low’ for inherent risk.

A review of the results highlights several issues common to archival research. First, by allowing the participating offices to select the subject client, final responses were not a random sample and were heavily weighted toward one industry, thereby reducing or eliminating between-industry analysis. Second, once survey forms were received, each open-ended question had to be interpreted and coded for analysis by the researcher. Such interpretation and related measurement issues may introduce researcher bias into the data being examined. Third, definitional issues may be introduced because of between-auditor differences in interpretation or inappropriate conclusions drawn by the researcher.

Bedard notes other limitations to the research. Since some questions involve judgments by the senior and the survey was removed in time from the audit, the quality of those judgments may vary or be biased. Also, since firm records and personnel are relied on, there may be cases in which workpapers, or the auditors’ responses, do not reflect complete or accurate information about the audit. Finally, because only three CPA firms participated and the sample clients were restricted as to industry, the ability to generalize the findings is limited.

Mock and Wright (1993)

Mock and Wright (1993) used a random sample of 600 engagements from the national client list of one CPA firm, with equal representation of manufacturing and merchandising companies. A data-gathering instrument requesting data from the two most recent audits was sent to audit partners near the end of the most recent audit. This timing was intended to promote accurate responses and minimize completion time, since most of the information requested came directly from current and prior-year working papers.

From the 600 companies randomly selected, 255 were eliminated because they did not meet certain specified criteria. Of the 345 remaining engagements, only 46 percent of the partners responded, giving a final sample of 159 audits: 84 manufacturing, 71 merchandising, and four not specified.

To address the research questions concerning the extent that audit plans are risk-adjusted, data was sought on the nature and extent of

Table 2: Illustrations of Archival Audit Research

<table>
<thead>
<tr>
<th>Archival Study Topic</th>
<th>Illustrative Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosticity of Audit Procedures; Error</td>
<td>Hylas &amp; Ashton (1982)</td>
</tr>
<tr>
<td>Characteristics of Audit Populations</td>
<td>Kreutzfeldt &amp; Wallace (1986)</td>
</tr>
<tr>
<td></td>
<td>Wright &amp; Ashton (1989)</td>
</tr>
<tr>
<td>Internal Control Evaluation &amp; Documentation</td>
<td>Mock &amp; Willingham (1983)</td>
</tr>
<tr>
<td>Audit Program Planning</td>
<td>Bedard (1989)</td>
</tr>
<tr>
<td></td>
<td>Mock &amp; Wright (1993)</td>
</tr>
<tr>
<td></td>
<td>Quadackers et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>Mock &amp; Wright (1999)</td>
</tr>
<tr>
<td>Inherent &amp; Control Risk Assessments</td>
<td>Waller (1993)</td>
</tr>
<tr>
<td>Audit Quality</td>
<td>Panel on Audit Effectiveness (POB 2000)</td>
</tr>
<tr>
<td>Analytical Procedures Effectiveness</td>
<td>Eimers (2001)</td>
</tr>
<tr>
<td>Audit Pricing and Planning</td>
<td>Johnstone &amp; Bedard (2001)</td>
</tr>
<tr>
<td>Fraud Risk Assessments</td>
<td>Mock &amp; Turner (2001)</td>
</tr>
</tbody>
</table>
audit procedures (planned and actual) for the two years, as well as assessments of inherent and control risks. A test instrument was developed and pilot tested to abstract the data from actual working papers, and once finalized, was sent to the engagement partner accompanied by a letter from the firm’s executive office requesting participation in the project. All responses were anonymous and returned by mail. If present, client names were redacted by the audit staff to maintain confidentiality. The senior on the audit was requested to consult the working papers to complete the instrument, and both the manager and the partner were asked to review all responses for thoroughness and accuracy.

To complete the task, the audit senior was asked to do the following:

1. Submit a copy of the audit program for both years.
2. Provide an evaluation of macro risk factors for the client. Macro risk factors were identified on a firm-standardized risk evaluation form. All risk factors were measured on a six-point, equal-interval scale with verbal end-anchors. Additional questions such as degree of complexity required additional non-documented judgments.
3. Provide an evaluation of control risk factors by account. Control risk for each account was identified on a firm-standardized control evaluation form. Risk was measured in terms of degree of reliance on controls: strong controls, moderate controls, weak controls, and not relied on.
4. One of the final two sections of the test instrument requested rationale statements briefly explaining the reason(s) for any significant changes made in the nature, extent, or timing of procedures between years. The last section collected demographic information about the individual completing the instrument.

This form of data-gathering was designed to reduce subjective judgments in the measurement process. As noted by Mock and Wright, however, the coding of open-ended audit programs by the researchers can be difficult because of technical terminology and/or because individuals may describe the same audit test in different or imprecise ways. For those audit programs containing non-standard procedures, one of the two researchers and an independent person performed a content analysis with an initial agreement level of 78 percent. Similar procedures were required in coding the rationale memos. Another issue noted by Mock and Wright is that the gathering of data for both years concurrently creates the potential for non-independent risk assessments or confounding factors such as a change in the audit team or firm technology. Additionally, changes in the overall audit environment and in professional standards may reduce the generalizability of results and introduce confounding factors that may have influenced observations.

Mock and Turner (2001)

In a study sponsored by the AICPA Auditing Standards Board (ASB), Mock and Turner (2001) used an archival approach to examine actual fraud risk assessments and their effects on audit programs following the issuance of SAS No. 82 (AICPA 1997). With assistance from representatives of the AICPA and the ASB, three national CPA firms were contacted and agreed to participate in the research project. After identifying liaisons within each firm, copies of each firm’s documentation regarding integrating fraud risk assessments into the planning process were obtained.

Documentation obtained from each CPA firm disclosed that the firms had responded to SAS No. 82 in different ways. While each firm appeared to comply with the requirements of the SAS, the different approaches made development of the survey instrument difficult. To be useful, the survey instrument requested information in such a way that the data returned could be combined for analysis, but designed to allow completion in an economically acceptable time period.

One firm indicated concerns regarding client-confidentiality and for that firm, a survey instrument was designed requiring firm personnel to transcribe information from the audit workpapers to the survey instrument. The concerns of the other two firms related primarily to the cost involved in completing the survey and for those firms, a survey instrument was designed asking each respondent to attach relevant redacted fraud assessment workpapers and complete only a short set of questions. Extensive pilot tests of each form of survey instrument identified several issues, but only
minor modifications to instrument design were required.

Final survey instruments were provided to eighty partners in each of the three participating firms. Each partner was asked to complete the instrument for two clients for a two-year period — one identified as being ‘low risk’ and one considered having a risk ‘other-than-low.’ Risk categories were based on individual firm assessment methods rather than being specified by the researchers and did not necessarily refer to the risk of fraud. Thus, some ambiguity was introduced based on differences in terminology used by each CPA firm. Additionally, for each ‘other-than-low’ risk client selected, partners were asked to select a low-risk firm, matched, if possible, within the same industry and of similar size. Timing also was an issue, as one firm provided fraud risk-assessments for 1998 and 1999 while the other two firms provided data for 1997 and 1998.

The final phase of the data-gathering process occurred over a twelve-month period and resulted in a total of 101 useable responses (42 percent response rate.) These responses provided data for 202 different companies or 404 company-years.

Because of widely differing terminology and approaches to the assessment of fraud risk, coding of data was difficult. To compensate for such differences, the researchers used the framework provided by SAS No. 82 to categorize fraud risk factors. This approach was made more complex because of within-firm changes over time in fraud risk-assessment procedures. Even with this generalization, however, substantial differences between firms resulted in data being aggregated to a greater extent than might be desired. Additionally, for one firm, audit team members were required to examine the audit workpapers and to interpret the contents in the context of the survey instrument. For the other two firms, the researchers themselves had to examine and interpret workpaper copies. Results from this study were presented to the Auditing Standards Board for analysis of the need to modify SAS No. 82.

In addition to these three examples of archival audit research, one other recent major study used archival methods to gather data and should be noted. At the request of the chairman of the U.S. Securities and Exchange Commission (SEC), in October 1998, the Public Oversight Board\(^\text{14}\) (POB) appointed the Panel on Audit Effectiveness to assess whether independent audits adequately serve and protect the interests of investors. As part of this study, the Panel assessed recent trends in audit practices to determine whether they are in the public interest. To accomplish this, the Panel performed Quasi-Peer Reviews (QPRs) of the quality of 126 audits of SEC registrants in 28 offices of the eight largest audit firms. Each QPR focused on audit methodologies, policies and procedures manuals and other guidance materials, information about risk management and professional development activities, audit-related marketing publications, and policies and procedures for recruiting, evaluating, compensating and promoting audit personnel. Based on the QPRs, in August 2000, the Panel made numerous recommendations directed to the U.S. Auditing Standards Board, the audit firms, the SEC Practice Section of the AICPA, and to the audit committees of SEC registrants. To date, no subsequent studies have evaluated how these entities responded to the recommendations.

Unfortunately, details of the QPRs were not made public and all archival data were destroyed shortly after the report was made public. Thus the scientific value of this very promising archival study was severely compromised. Destroying data makes it impossible to determine how the various issues related to archival research were resolved. It also severely limits the ability of the scientific community to assess either the internal or external validity of reported findings.

**RECOMMENDATIONS**

There are many possible avenues to examine in making archival research more feasible, both for researchers and for audit firms. We will discuss several options that might provide favorable benefits.

First, the profession should study the possibility of creating a ‘safe harbor’ for audit firms to protect them from inadvertent legal or ethical violations while participating with independent researchers. Creation of such a safe harbor may well encourage audit firms to allow researchers greater access to archival data.

Second, a central clearing process might be considered. Presently, unless a study is commissioned by a professional body or by a specific
CPA firm, individual researchers must approach audit firms with proposal ideas. Because of the number of proposals received and substantial variation in the nature and quality of the proposals, firms may find it easier to simply reject most or all. If a central committee existed consisting of several academic researchers and a representative from each of the large firms, the process might be more effective. In this scenario, academic committee members would evaluate research proposals for the potential contribution to the literature and for the validity of the research design. Practitioners would review the proposal to evaluate the potential contribution to audit practice. One or more firms then could select specific proposals with which to participate and would be assured that the research design was adequate.

A third approach that might be used in conjunction with a central clearing committee is designing collaborative projects where academic researchers are teamed with one or more practitioners. Several archival studies have taken this approach, including Hylas and Ashton (1982); Mock and Willingham (1983); Kreutzfeldt and Wallace (1986); and Quadackers et. al. (1996). Eimers (2001) discusses several advantages of collaborative archival research including improved data quality due to direct researcher access to the participants. Most importantly, direct contact with participants enables researchers to request additional clarifications when specific replies are not clear.

In many countries such as The Netherlands, a significant number of audit professors also are active practitioners. This tradition provides greater opportunity for the academic researcher to obtain direct access to archival data and could be considered a form of collaboration.

An enhancement to collaborative projects is allowing academic researchers to participate in internships with an audit firm. Such an arrangement can free the academic from teaching and administrative assignments as well as provide the benefits noted by Eimers, above. Again, if a central clearing committee existed, applications for and availability of internships could be coordinated through the committee.

Fourth, reviews by standard setters of additional required disclosures should continue. For example, a recent requirement by the SEC for public companies to disclose audit, consulting and other fees has stimulated a wide range of new audit research. Several of these studies have received attention in publications such as The Wall Street Journal, indicating that given the right topics, the public may be receptive to academic research.

Archival data extracted from audit workpapers offer many opportunities for researchers. Archival research can provide reliable research information for the profession in a number of areas and provides an opportunity to assess issues such as the impact of audit guidance (e.g. Mock & Willingham (1983)), the impact of regulation (e.g. Mock & Turner (2001)), and audit quality (POB (2000)). The potential payoff for the profession in adding to our knowledge and the opportunity cost of not conducting this type of research is substantial and should not be dismissed.

ACKNOWLEDGEMENTS

The authors would like to thank Ken Trotman, Arnie Wright, Jean Bedard, Ganesh Krishnamoorthy, Gary Monroe and Hai Lu for their helpful suggestions.

NOTES

1. In-depth case studies (e.g. Erickson et al. (2000)) offer another opportunity to obtain data based on actual audits. However, to date there have been few published audit case studies, possibly because of the confidentiality issue faced in all field-based audit research.

2. A potential weakness, however, is that workpapers sometimes may not accurately reflect circumstances. For example, actual audit hours might not be reported if the staff was under pressure to meet budgets.

3. A discussion of limitations on generalizability follows in General Design Issues.


5. There is the important issue of who should abstract the data and whether that person is taking the task seriously. The latter concern led Mock and Wright (1993, 1999) to require the audit manager or partner to ‘sign off’ that the instrument was complete and accurate.

6. To ensure attention to the research task, Gooding and Monroe (1998) paid the audit firm at commercial rates to collect data.
Audit staff were informed that time spent in completing a questionnaire would be billed and were instructed to include the time required to complete the questionnaire on their time sheets as billable hours.

7. After reduction for non-useable responses.
8. The study was an internal Arthur Andersen project, with compulsory participation.
9. Based on a random sample of 600 clients, with 255 engagements eliminated that did not meet the selection criteria or were no longer a client.

10. For example, Amer et al. (1994) examine between-auditor variance in the numerical interpretation of probability phrases used in audit settings, as well as on the degree to which auditors are aware of the variance. Results reveal substantial between-auditor variance in interpretation and generally low levels of variance awareness.

11. The authors are not aware of any lawsuits where information from academic researchers has been introduced into court to the detriment of a CPA firm. There is no question however, that such potential exists. Also there is the concern that if a number of studies suggest weak auditor performance, regulators may feel the need to change laws or regulations in response.

12. A fourth firm initially agreed to participate but withdrew prior to pilot testing the survey instrument.

13. In addition to documentation differences between firms, within-firm differences also were noted as each firm revised its fraud assessment approach based on experience. Also, three of the firms were revising their overall audit approaches, thereby exacerbating the difficulties in obtaining comparative data over time.

14. The POB was created in 1977 as an independent private sector body charged with overseeing and reporting on the programs of the SEC Practice Section of the AICPA.

REFERENCES


**AUTHOR PROFILES**

**Theodore J. Mock** is the Arthur Andersen Alumni Professor of Accounting at the University of Southern California and professor of auditing research at Maastricht University in the Netherlands. In 1983 he founded the USC Audit Judgment Symposium (now the International Symposium on Audit Research). Ted has served many positions within the American Accounting Association including editor of Auditing: A Journal of Practice & Theory, Director of Research, and President of the Auditing Section.

**Jerry L. Turner** is an assistant professor of accounting at The University of Memphis in Memphis, Tennessee. His primary teaching and research interests are in auditing and accounting information systems. In addition to two research monographs, he has published in the Journal of Accounting, Auditing and Finance, the Journal of Forensic Accounting, and other journals.