

A Comparative Analysis Of Audit Service Supply Using Desk And Working Paper Reviews

Donald R. Deis Jr., University of Missouri-Columbia
Anna M. Rose, (Email: aniarose@montana.edu), Montana State University

Abstract

The Texas Education Agency (TEA) performs extensive reviews of audit working papers for selected Texas school district audits to evaluate the quality of audit services rendered by independent accountants. These reviews are referred to as Quality Control Reviews (QCRs). The TEA also conducts desk reviews of all audited financial statements issued by Texas independent school districts. During the desk review process, the TEA performs a limited check of the accuracy of financial statement amounts, completeness of financial disclosures, and existence of proper audit output. Prior research suggests that desk reviews and quality control reviews both measure audit quality (Colbert and O'Keefe 1995; Copley et al. 1994; Deis and Giroux 1992, 1996; Giroux et al. 1995; O'Keefe and Westort 1992; O'Keefe et al. 1994). The purpose of the paper is to determine if desk reviews capture the same audit quality information as quality control reviews. If so, regulators should focus resources on desk reviews since they are more timely and economical than QCRs. The results of the study, however, indicate that desk reviews and quality control reviews do not measure the same constructs. The findings suggest that desk reviews measure the industry specific knowledge of the auditor, which is only one aspect of audit quality. Moreover, it was observed that some audits passing the desk review were graded low in quality by the working paper review. This finding suggests that low audit quality may be more prevalent than generally suspected and that steps to improve audit quality and auditor credibility are warranted.

Introduction

A string of recent accounting failures (e.g., Enron, WorldCom, Cendant: Waste Management) has increased the scrutiny of public accounting firms. On December 2, 2001, Enron filed for Chapter 11 bankruptcy without any warnings from its Big 5 auditor, Arthur Andersen. In another failure, Cendant's shareholders filed a lawsuit accusing its independent auditor of providing a "clean" audit opinion without gathering adequate documentation and without reviewing company general ledgers (Pachelle and MacDonald 1999). Such instances shake financial statement user faith in the credibility of the auditor and the accounting profession. Based on recent auditing debacles, it is evident that substandard audit reports can be prepared without performing necessary audit procedures. In fact, various audit resources (e.g., sample reports, industry guidelines, reporting software) make it possible to conduct an audit and prepare an audit report with minimal effort. In the private sector, the need for independent measures of audit quality is evident.

In an effort to improve audit quality, perhaps the profession can look to solutions developed in the governmental auditing arena. This study investigates two specific audit quality measures used to evaluate the audits of independent school districts. Regulators have used two types of reviews to assess the auditor's compliance with GAAS and GAGAS: (1) desk reviews and (2) working paper reviews. Since the two reviews differ in scope and purpose, their relationship and appropriate use is unclear. The purpose of this study is to determine if desk reviews capture the same audit quality information as working paper reviews.

Extant research concentrates on the following three indicators of audit service supply: (1) audit hours, (2) compliance with generally accepted reporting standards as evidenced in desk reviews of audit reports, and (3) compliance with generally accepted auditing standards as determined in working paper reviews (i.e., quality control reviews or “QCRs”). None of these measures is a perfect indicator of the quality of audit supply, and each may capture unique aspects of the supply of audit services. Audit hours can, for example, be considerable in absolute terms but can be used inefficiently. In addition, required procedures may be omitted, and the audit reports can be incorrectly prepared. Following GAAS reporting guidelines results in a “clean” desk review, but meeting reporting requirements can be accomplished even if the underlying audit work is substandard. Finally, collections of fieldwork evidence (i.e., working papers) do not always translate into high quality audit reports. The quality of audit reports is also influenced by the methods used to synthesize evidence and independence issues (e.g., the auditor’s willingness to report sensitive matters). To develop a better view of audit service supply, this paper considers all three audit service supply metrics in unison.

The data used in this study contains 178 audits of independent school districts in Texas on which state auditors in the oversight agency conducted both a desk review and QCR. Actual total audit hours for each audit were also available. Data analysis revealed that reports meeting generally accepted reporting standards sometimes mask low quality audit work and that, occasionally, audits that collected adequate audit evidence were found to have reporting violations. Such flaws in the audit process point to potential threats to auditor credibility. Results also indicate that desk reviews and quality control reviews do not measure the same aspects of audit service supply. Desk reviews appeared to capture the industry-specific knowledge of the auditor. All findings were confirmed using both ordinary least squares regression and simultaneous equations analysis.

Background

As a result of the Single Audit Act of 1984, the Department of Education now monitors the quality of audits of public school districts. In Texas, the authority to review audits of independent school districts (ISDs) has been delegated to the Texas Education Agency (TEA). The regulatory actions of the TEA have provided researchers with an independent assessment of audit quality. The Texas Education Agency (TEA) performs extensive reviews of audit working papers for selected Texas school district audits in order to measure audit quality. These reviews are referred to as Quality Control Reviews (QCRs). The TEA also conducts desk reviews of all audited financial statements issued by Texas independent school districts. During the desk review process, the TEA performs a limited check of the accuracy of financial statement amounts, completeness of financial disclosures, and existence of proper audit output.

In the private sector, proxies for audit quality are difficult to construct. The surrogate used in most prior research has been audit firm size. According to DeAngelo (1981b), large auditors perform better quality audits because they have many clients and would lose more quasi-rents than smaller firms if they provided poor quality audits. In addition, large auditors are more independent from their clients because each client’s fee is only a small percentage of the auditor’s total revenues. In the public sector, the TEA develops a reliable ex-post measure of audit quality for school districts.

Several studies have used state agency data to examine the determinants of audit quality, audit fees and audit hours for public school district audits. The results have been consistent between the two different audit quality measures that have been developed to date. The three Texas school district studies (Deis and Giroux 1992, 1996; Giroux et al. 1995) used an audit quality metric developed from quality control review (QCR) letters prepared by the TEA. O’Keefe et al. (1994) used desk reviews prepared by the California State Controller’s Office to create the second quality measure.¹ The current study uses both desk review and working paper review (QCR) data to more thoroughly examine the supply of audit services. The purpose of the paper is to determine if desk reviews capture the same audit quality information as QCR.

An important question for regulators is whether desk reviews and QCRs are redundant measures of audit quality. Many audits in the public sector are of low quality and do not conform to GAAS (US GAO 1986). As a

result, it is essential that some formal review process be conducted to promote quality financial reporting by governmental units. The TEA could, however, significantly reduce public spending if performing only desk reviews would accomplish the monitoring goals set by the Department of Education. On the other hand, if desk reviews measure different aspects of quality than QCRs, it may be necessary to retain both desk reviews and QCRs, or even expand the QCR program.

This research uses the audit quality and audit hours models developed in Deis and Giroux (1992, 1996) and Giroux et al. (1995). Also, the database used in the above studies is further developed with the addition of a new metric for desk reviews. The findings of O'Keefe et al. (1994) (the only other study with detailed desk review data), suggest that the desk score variable will yield results very similar to those found for QCR score in Deis and Giroux (1996) and Giroux et al. (1995). Consistent results for desk reviews and QCRs would indicate that both scores measure very similar audit quality characteristics. One of the key advantages of the current study is that the database allows for a direct comparison of desk reviews, QCRs, and audit hours.

The remainder of the paper is organized as follows. The next section presents a discussion of prior research on school district audits. The following sections discuss the models and methods used in the study. Last, the results of tests performed and concluding remarks are presented.

Prior Research

In the late 1970's, many cities experienced financial difficulties coupled with inadequate financial records. The unexpected crisis in New York City and other cities led to passage of the Single Audit Act of 1984 in order to improve the quality of audits and financial reporting of governmental units. Among other provisions, the Single Audit Act requires desk reviews and quality control reviews for Independent School Districts (ISD) (Deis et al. 1990). These reviews have been used to monitor audit quality. At the Texas State level, the Department of Education has delegated reviewing authority to the Texas Education Agency (TEA). Desk reviews performed by the TEA are preliminary reviews of all audited financial statements issued by Texas ISDs. In the desk review, TEA auditors check that certain financial statement amounts are accurate, that financial disclosures are complete, that the report was filed on a timely basis, and that correct opinion format was used.

Problems that are discovered in the desk review phase could cause the TEA to perform a quality control review of the auditor. In such a review, the agency examines the auditor's working papers. To date, QCRs are performed for only a limited number of audits, although the TEA intends to expand the program to include all audits (Deis et al. 1990). Auditors may be chosen for a QCR on a regional basis or as a result of a desk review. QCRs are much more comprehensive than desk reviews and focus on evaluating whether auditor's working papers indicate the proper application of GAAS and GAGAS procedures. When QCRs suggest that audits are of particularly low quality, the TEA can refer the auditor to the Texas State Board of Accountancy for disciplinary action (Deis et al. 1990).

Previous studies have primarily used the results of quality control reviews to test audit quality theories. Using TEA data, three studies analyzed relationships between audit quality, audit fees and audit hours. Deis and Giroux (1992) developed a quality metric for 232 ISD audits from QCR letters of findings sent to the auditors. Prior research of commercial firms has shown that reputation differences and power conflicts are valid explanations for differences in audit quality between small and large auditors (see e.g., DeAngelo 1981b; Nichols and Price 1976; Rubin 1988). The purpose of the Deis and Giroux study was to determine if these factors explained variations in audit quality for a homogeneous group of auditors (i.e., small auditors of Texas ISDs). Panel A of Table 1 summarizes the results of their study.

The following variables were used in the three Texas school district studies. These variables deserve some explanation because they are also included in the present study. All of the variables described below are statistically significant in Deis and Giroux (1992). In the model, increases in auditor tenure result in decreases in audit quality, a result consistent with private sector research (see e.g., Francis and Simon 1987; Turpen 1990). This finding is also

consistent with DeAngelo's (1981b) quasi-rent theory, which states that an auditor decreases audit quality over time (by not reporting all deficiencies) in order to retain the client and receive future fees. Audit quality was found to increase with increases in the number of ISD clients of an auditor (see also, Shockley 1982). An auditor with many ISD clients has more industry expertise and does not want to damage its reputation with poor quality audits.

The variable PEER measures whether the auditor is a member of the AICPA Peer Review Section. Auditors that are members provide higher quality audits because they commit to higher professional standards and can better resist client pressures. In a later study, Giroux et al. (1995) documented that participation in peer reviews was also associated with increased audit fees. It appears that firms that voluntarily join peer review programs are signaling commitment to the profession. They provide higher quality audits for higher fees. The study did not find an association between peer review membership and audit hours.

As client size and client wealth increase, audit quality decreases. The power-conflict explanation of this relationship is that larger, wealthier clients can more easily influence audit results. TIME measures when the TEA receives an audit report. Less timely and late reports are associated with lower quality. The last significant variable in the model is HOURS (the actual hours spent to perform an audit). As hours increase, audit quality increases. This is consistent with the proposition that auditors expend more effort on higher quality audits. Deis and Giroux interpret these results based on theories of reputation and power-conflict. Their study's results confirm that both of these explanations for differing audit qualities for small versus large audit firms are also valid for a homogeneous group of small auditors.

Deis and Giroux (1996) further analyzed the TEA data. The purpose of the study was to determine what effect initial audits had on audit quality, audit fees and audit hours (see Table 1, Panels A and B). Most variables in the quality model are the same as those in Deis and Giroux (1992). A variable was added (YEAR 1) to distinguish between initial and subsequent audits. The study found a significant relationship between YEAR1 and QCR scores, where initial engagements produced higher audit quality. Initial engagements were also associated with lower fees and higher audit hours (see Table 1, Panel B). These results support DeAngelo's (1981a) low-balling hypothesis. Auditors price initial audits at a discount in order to capture the future profits associated with new clients - even though more hours are required to perform first year audits (Francis and Simon 1987; Simon and Francis 1988; Turpen 1990). Once a client chooses an auditor, the costs associated with switching to a new auditor are very high. In addition, the auditor will earn profits in future years because less audit hours will be required as a result of the learning curve. Auditors provide high quality for initial audits because they do not want to jeopardize this future profit potential. The results of Deis and Giroux extend private sector low-balling theories to the public auditing sector.

A study by O'Keefe, King and Gaver (1994) used desk review results to test audit quality theories very similar to the three TEA studies mentioned above. The quality metric was based on desk reviews performed by the California State Controller's Office. The controller's office reviews the financial statements and audit reports of school districts to check compliance with GAAS and GAAP standards. Review results are sent to the school districts and the auditors. The study analyzed reviews of 935 reports for the year 1986. The purpose of the study was to determine if compliance with GAAS increases with increased labor (where audit fees were used as a proxy for labor) and with increased industry specialization of the auditor (number of ISD clients). Panel A of Table 1 compares the desk review results with the results of the TEA studies. Consistent with Deis and Giroux (1992, 1996), O'Keefe et al. found that an increase in the number of school district audits decreased the number of violations of GAAS and, therefore, increased audit quality. Similarly, fewer violations were associated with higher audit fees. Audit fees proxy for the amount of labor and effort expended to perform the audit. Audit quality should improve as effort increases because of the greater likelihood of identifying material misstatements in the financial statements and of properly following GAAS regulations. In general, the results of O'Keefe et al. were consistent with the TEA studies.

Table 1: Summary Results Of Audit Quality And Fee Studies For Independent School Districts PROPOSITIONS

Panel A

AUDIT QUALITY MODELS

Deis and Giroux 1992		Deis and Giroux 1996		O'Keefe et al. 1994		Variables
Effect on Quality		Effect on Quality		Effect on Quality		
as Variable		as Variable		as Variable		in all models
Increases		Increases		Increases		proxy for:
TENURE	decreases	YEAR1	increases	CPA_CHANGE	increases	initial audit
CLIENTS	increases	CLIENTS	increases	#SD_AUDITS	increases	industry knowledge
PEER	increases	PEER	increases	CPA_SOCIETY	increases	peer review
SIZE	decreases	SIZE	decreases	CLIENT_REVENUE	decreases	size of ISD
WEALTH	decreases	WEALTH	decreases			
TIME	decreases	TIME	decreases			
HOURS	increases	HOURS	increases	AUDIT FEE	increases	labor
		OFFICES	increases	BIG8	increases	brand name

Variables listed above are statistically significant variables.

Full models:

Deis and Giroux 1992

$$\text{Ln}(\text{QUALITY}) = \text{TENURE} + \text{CLIENTS} + \text{PEER} + \text{BOARD} + \text{SIZE} + \text{WEALTH} + \text{YEAR} + \text{REPORT} + \text{TIME} + \text{HOURS}$$

Deis and Giroux 1996

$$\text{Ln}(\text{QUALITY}) = \text{YEAR1} + \text{YEAR2} + \text{CLIENTS} + \text{PEER} + \text{SIZE} + \text{WEALTH} + \text{TIME} + \text{HOURS} + \text{OFFICES}$$

O'Keefe et al. 1994

$$\text{VIOLATIONS} = \text{TIER2} + \text{BIG8} + \text{CPA_CHANGE} + \text{\#SD_AUDITS} + \text{CPA_SOCIETY} + \text{DIVISION_FIRM} + \text{CLIENT_REVENUE} + \text{AUDIT_FEE} + \text{PREV_PROBLEMS} + \text{WEAK_CONTROLS}$$

Panel B

AUDIT FEE AND HOUR MODELS

Deis and Giroux 1996

	Effect on <u>Fees</u>	Effect on <u>Hours</u>
	as Variable	as Variable
	Increases	Increases
YEAR1	decreases	increases
CLIENTS	decreases	decreases
SIZE	increases	increases
OFFICES	increases	increases
CAFR	increases	increases
ICREPORT	increases	increases
OPIN	increases	increases
PCI	decreases	decreases
QUALITY	increases	increases

Variables listed above are statistically significant variables.

Full model:

Deis and Giroux 1996

$$\text{FEE or HOURS} = \text{YEAR1} + \text{YEAR2} + \text{CLIENTS} + \text{SIZE} + \text{OFFICES} + \text{CAFR} + \text{ICREPORTS} + \text{OPIN} + \text{PCI} + \text{QUALITY}$$

The results of O’Keefe et al. (1994) begin to suggest that desk reviews capture much of the same audit quality information as quality control reviews. If desk reviews did measure the same aspects of audit quality as QCRs, then the quality control review process would be redundant, and considerable public funds could be saved by eliminating the QCR process. The question is particularly relevant given that the TEA is considering the implementation of QCRs for every Texas school district.

The problem with basing any policy decisions on the O’Keefe et al. (1994) study is that no direct comparison was made between QCRs and desk reviews. The study found that desk review scores were positively associated with audit fees and the number of ISD clients held by the auditor. These results do not offer convincing evidence that desk reviews capture full audit quality information. For example, if the desk review score captured information about the industry expertise of the auditor, we would expect to find very similar results. More expertise would be associated with auditors that have many school district clients, and more expertise would result in a demand by the auditor for higher fees.

We propose that desk reviews do not capture the same audit quality information as QCRs. Desk reviews are based on the content and format of the audited financial statements, and no analysis of audit working papers is involved (TEA 1994a). The desk review process may produce information relating to the skill of the auditor, which is associated with audit quality, but will not reveal the quality of the audit itself.

Proposition 1. The determinants of desk reviews will not be the same as described for quality control reviews by Deis and Giroux (1996) and Giroux et al. (1995).

Proposition 2. The desk review score will be a reliable predictor of QCR score.

O’Keefe et al. used audit fees as a proxy for measuring audit labor. They found that increased quality, as measured by desk review score, was associated with higher audit fees. This association supported their prediction that desk reviews measured audit quality. In this study, we use a direct measure of audit labor – the actual number of hours spent on the audit by the auditor. We propose that the desk review score is not capturing the same audit quality information as a quality control review. Rather, desk review scores likely measure some aspect of auditor expertise.

As a result, desk review scores will not be related to the level of labor input to the audit process, as measured by actual audit hours.

Proposition 3. Desk review scores will not have a positive and significant association with audit hours.

Models and Methods

TEA response letters written to ISDs were analyzed to develop a desk review metric. The response letters describe the results of the individual desk reviews. Table 2 shows the categories of the deficiencies reported in desk review letters and the frequency count associated with each deficiency. Deficiencies include errors in financial statements and auditor related errors. Errors in financial statements addressed by the TEA in the review letters included only the errors not reported in the audit reports. We located 181 of the 232 audit reports used in the Deis and Giroux (1996) and Giroux et al. (1995) studies and categorized each deficiency in the 181 desk review letters. A detailed description of the types of violations included under each category can be requested from the authors.

Table 2: Desk Review Scoring

Desk Review Score Items	Deficiency Severity	Assigned Weight	Frequency	Percent*
Late filing of audit report	Major	2	7	1.27%
Certificate of Board	Minor	1	8	1.45%
Errors in auditor-supplied information				
Missing required disclosure or schedule	Major	2	47	8.53%
Auditor-specific errors	Minor	1	4	0.73%
Budget				
Legal procedures not followed	Minor	1	33	5.99%
Incorrect reporting	Major	2	66	11.98%
Financial Statements				
Incorrect classifications	Major	2	42	7.62%
Disagreement between schedules and F/S	Minor	1	44	7.99%
Incomplete or missing disclosures	Minor	1	4	0.73%
Incorrect reporting procedures	Major	2	3	0.54%
Legal compliance problems	Major	2	117	21.23%
Notes to F/S				
Missing F/S notes	Major	2	21	3.81%
Missing disclosures within reported notes	Minor	1	24	4.36%
Disagreement between notes and F/S	Minor	1	15	2.72%
Special Schedules				
Incorrect classifications	Minor	1	1	0.18%
Incomplete or missing disclosures	Minor	1	9	1.63%
Incorrect reporting procedures	Minor	1	3	0.54%
Missing F/S	Extreme	3	9	1.63%
Missing schedules	Major	2	20	3.63%
Improper format in schedule or statement	Minor	1	17	3.09%
Math errors	Minor	1	38	6.90%
Prior years' deficiencies repeated in current year	Major	2	19	3.45%
<i>Total number of deficiencies found in the 181 reports</i>			551	
*Percent of the total number of deficiencies (551)				

The categories of deficiencies in Table 2 are classified into minor, major and extreme deficiencies. Classification of deficiencies was based upon interviews with the TEA, analysis of the TEA desk review checklist and audit review manual, and examination of the TEA criteria for the referral of desk reviews for quality control reviews (TEA 1994a,b). The extreme deficiency category includes items that would automatically cause the referral of an audit for a QCR. The extreme deficiency category is weighted three times more severe than the minor deficiency category due to the TEA's perceived gravity of the violations.² Major deficiencies are assigned two times more weight than minor deficiencies. Major deficiencies were those identified by the TEA that would warrant referral for a QCR if multiple violations were present. The desk review metric is initiated by summing the weighed deficiencies for each school district multiplied by minus 1. The resulting index ranged from 0 to minus 17. The dependent variable DESKSCORE was constructed by adding 18 to the desk review index and dividing the result by 18 as follows:

$$\text{DESKSCORE} = \frac{(\text{Desk Review Weighted Index} + 18)}{18}$$

The result is a variable with values from 0 to 1. Higher (lower) DESKSCORE values indicate less (more) violations detected in the desk review process.

The procedure to construct the metric for audit service quality found in the QCR was similar to that used to develop DESKREVIEW. As described by Deis and Giroux (1992, 468-469), 19 categories of deficiencies in the audit were drawn from TEA's letters of findings (see Table 3). TEA's Director of Audits ranked the deficiencies in order of importance to the decision to refer the audit to the Texas State Board of Public Accountancy for possible sanctions. To construct the metric in this study, the most important items were ranked minus 19 and the least important was weighed minus 1. By summing these items for each audit, a weighed index of QCR findings was constructed. This index ranged from minus 1 to minus 131. The dependent variable, QCRSCORE, was constructed by adding the weighed QCR index to 132 and then dividing the result by 132 as follows:

$$\text{QCRSCORE} = \frac{(\text{QCR Weighted Index} + 132)}{132}$$

The resulting metric ranges between 0 and 1. High (low) values for the QCRSCORE indicate high (low) levels of violations found in the working paper review.

One advantage of the two audit quality metrics is that they are both on a scale ranging from 0 to 1. Moreover, they can be interpreted as probability ratios of the quality of audit service supplied. A plot of DESKSCORE and QCRSCORE appears in Figure 1. DESKSCORE is plotted on the vertical axis. QCRSCORE is plotted along the horizontal axis. The referral outcome was used as a symbol to plot each observation (1 = referred and 0 = not referred). The figure shows a number of interesting phenomena. First, a number of high DESKSCORE audits have a low QCRSCORE (upper left part of the figure). Most of these high DESKSCORE/low QCRSCORE observations were referred to the State Board for sanctions. This is an example of "window dressing" (reports in proper format but without adequate underlying audit evidence and documentation to support the audit report). Second, a number of low DESKSCORE audits have a high QCRSCORE (lower right part of the figure). None of these audits were referred to the State Board. It appears that the referral decision is influenced more by the results of the QCR than the desk review.

The three propositions stated earlier are investigated using two audit quality models and an audit hours model similar to those used in prior literature (Deis and Giroux 1992 and 1996; Giroux et al. 1995). Two types of analysis are conducted first using ordinary least-squares regression (OLS) and then using simultaneous equations analysis where the audit service supply and audit quality constructs are jointly estimated. The three OLS models used in this study follow.

Audit Quality Models

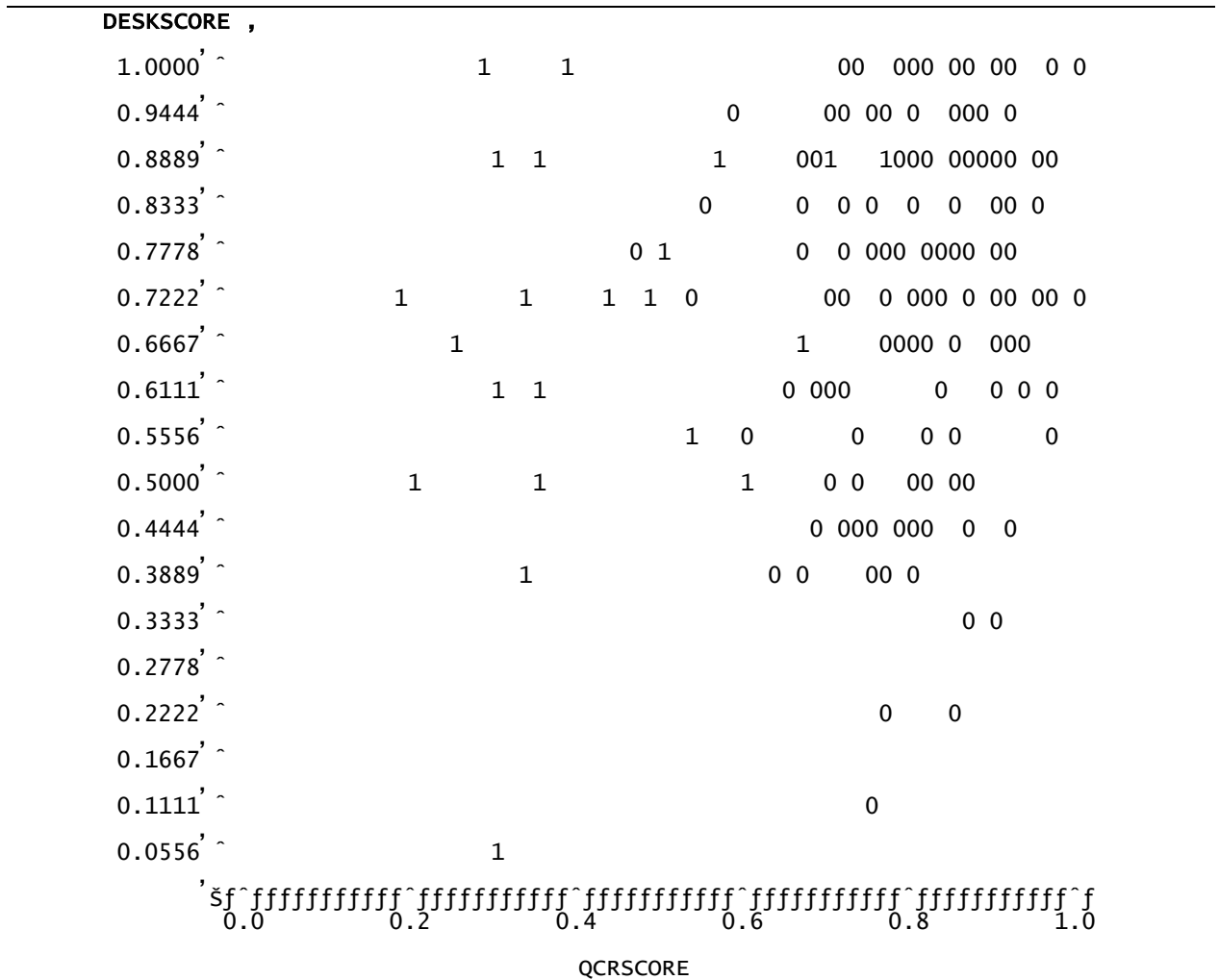
$$\text{DESKSCORE} = \alpha_0 + \beta_1\text{SIZE} + \beta_2\text{WEALTH} + \beta_3\text{CLIENTS} + \beta_4\text{OFFICES} + \beta_5\text{PEER} + \beta_6\text{TIME} + \beta_7\text{HOURS} + \beta_8\text{YEAR1} + \varepsilon_1$$

$$\text{QCRSCORE} = \alpha_0 + \beta_1\text{SIZE} + \beta_2\text{WEALTH} + \beta_3\text{CLIENTS} + \beta_4\text{OFFICES} + \beta_5\text{PEER} + \beta_6\text{TIME} + \beta_7\text{HOURS} + \beta_8\text{YEAR1} + \beta_9\text{DESKSCORE} + \varepsilon_2$$

Audit Hours Model

$$\text{HOURS} = \alpha_0 + \beta_1\text{SIZE} + \beta_2\text{PCI} + \beta_3\text{OFFICES} + \beta_4\text{CLIENTS} + \beta_5\text{CAFR} + \beta_6\text{ICREPORTS} + \beta_7\text{OPIN} + \beta_8\text{YEAR1} + \beta_9\text{DESKSCORE} + \beta_{10}\text{QCRSCORE} + \varepsilon_3$$

Figure 1 Plot of DESKSCORE and QCRSCORE



NOTE: 43 obs hidden. Audit firms referred to Texas State Board of Public Accountancy are indicated by “1” and “0” if audit was not referred.

The variables in the models have the following definitions.

DESKSCORE: desk review score ranging from 0 (low) to 1 (high)

QCRSCORE: metric for audit quality based on QCRs ranging from 0 (low) to 1 (high)

SIZE: log of average daily attendance for the ISD

WEALTH: log of property values per student

CLIENTS: number of school district clients of the auditor

OFFICES: number of offices owned by the audit firm

PEER: 1 if audit firm is a member of AICPA’s Peer Review Section, 0 otherwise

TIME: number of days from fiscal year end (8/31) until the date of the audit report divided by 120 (school districts have until 12/31, 120 days from 8/31, to submit the annual report)

HOURS: log of actual hours required in performing the audit

YEAR1: 1 if the audit is a first year audit of the client, 0 otherwise

PCI: log of per capita income in the school district

CAFR: 1 if audit report was a Comprehensive Annual Financial Report, 0 otherwise (i.e., General Purpose Financial Statements)

ICREPORT: 1 if auditor reported material weaknesses in internal controls, 0 otherwise

OPIN: 1 if a qualified independent auditor’s report is issued, 0 otherwise

Table 3 QCR Scoring

QCR Deficiency Item	Assigned Weight	Freq.	Percent*
Audit Program: Major	-19	17	2.31%
Working Papers: Major	-18	3	0.41%
Legal Compliance: Major	-17	45	6.12%
Audit Procedures: Major	-16	31	4.22%
Internal Control: Major	-15	12	1.63%
No Management Representation Letter	-14	25	3.40%
Substantive Tests: Major	-13	17	2.31%
Errors in Financial Statement Presentation	-12	63	8.57%
Ethics Violations	-11	2	0.27%
Legal Compliance: Minor	-10	45	6.12%
Errors In Audit Report	-9	20	2.72%
Inadequate Risk Assessment	-8	97	13.20%
Inadequate Statistical Sampling	-7	118	16.05%
No Engagement Letter	-6	27	3.67%
Audit Program: Minor	-5	12	1.63%
Audit Procedures: Minor	-4	105	14.29%
Working Papers: Minor	-3	56	7.62%
Internal Control: Minor	-2	28	3.81%
Substantive Tests: Minor	-1	12	1.63%
<i>Total number of deficiencies found in 178 reports</i>		735	
<i>*Percent of the total number of deficiencies (735)</i>			

The audit quality models and hours model are analyzed using OLS regression. This study uses the database used in Deis and Giroux (1992 & 1996) and Giroux et al. (1995) with the additional variable, DESKSCORE. Using the same database allows for direct comparison of QCR scores and desk review scores. The sample consists of 181 ISDs and audit reports for fiscal years 1986 and 1987. As in the previous studies, Big 8 and regional audit firms were not included in the sample to eliminate any “brand name” bias associated with these firms. Also, each school district is represented only once in the sample (Deis and Giroux, 1996).

Table 4 presents descriptive statistics of the key variables. The average desk review variable (DESKSCORE) is 0.72 with a range from 0.06 to 1.00. In comparison, the average QCR review variable (QCRSCORE) was 0.76 and ranged from 0.01 to 0.99. Both scales are similar with a “passing” average in the 70th percentile. The ranges are quite large for both, however, indicating a wide variety in performance. The initial sample included 181 school district desk reviews. Three observations resulted in extreme outliers in the regression models. These observations were removed from all analyses due to their potential influence on the regression estimations. The final sample consists of 178 observations.

Results

Table 5 presents the Pearson Correlation Coefficients for selected variables. The majority of the independent variables do not have strong correlations, with the exception of HOURS and SIZE (0.63). The DESKSCORE and the QCRSCORE variables have a correlation of only 0.16. Preliminary analyses appear to indicate that these two variables are proxies for different constructs.

To investigate the first proposition, which suggests that DESKSCORE captures different audit quality characteristics relative to QCRs, the original Deis and Giroux (1996) model with DESKSCORE as the dependent variable was utilized. Table 6 shows the results for the audit quality model.³ The model is statistically significant, and the R² is 0.110. WEALTH, OFFICES, PEER, TIME, and HOURS are not significant. YEAR1 is significant, but

has the opposite sign of YEAR1 in Deis and Giroux (1996, 76) (see Table 1). If the audit is an initial audit (YEAR1 = 1) then DESKSCORE decreases. This suggests that initial audits conform less to GAAS reporting standards than repeat engagements. One explanation for this result is that DESKSCORE measures a different characteristic of quality than the QCR, such as the industry and client specific knowledge of the auditor. Client specific knowledge would necessarily be lower for new clients. The significant relation between number of ISD clients and DESKSCORE also suggests that desk reviews capture the industry and client specific knowledge of the auditor.

The second proposition concerns the ability of DESKSCORE to predict audit quality as measured by QCRSCORE. To test for this relation, DESKSCORE was introduced as an additional independent variable to the Deis and Giroux (1996) model of the determinants of audit quality (see Table 7). DESKSCORE is a significant predictor of QCR score at the .01 level. Based on this result and the sign of YEAR1 in Table 6, it appears that DESKSCORE is measuring industry specific and/or client specific knowledge of the auditor, which would be positively related to audit quality. During initial audits, auditors know very little about the client (and possibly about the industry) which would result in a low DESKSCORE.

Audit hours increased with higher quality scores in Deis and Giroux (1996) and were predicted to increase with higher quality as measured with desk review scores by O'Keefe et al. (1994). We find in Table 6, however, that audit hours are not a positive and significant determinant of DESKSCORE.⁴ As auditors input more labor to the audit process, the desk review score does not improve. This indicates that the desk review score is determined by knowledge that is in place prior to the audit, and increased auditor effort does not affect the score. The finding supports the proposition that DESKSCORE measures only one aspect of audit quality, namely client/industry specific knowledge.

Table 4 Descriptive Statistics

Variable	Mean* (Frequency)	Range
SIZE (log of average daily attendance)	3,994	36 to 44,776
WEALTH (log of property tax base per student)	243,600	21,979 to 5,404,050
CLIENTS (number of ISD clients)	3.32	1 to 21
OFFICES (number of offices for audit firm)	1.30	1 to 8
PEER (1 = audit firm member of AICPA's Peer Review Program, 0 otherwise)	(0 = 151, 1 = 27)	0 to 1
TIME (portion of 120 days after yearend to file audit report with TEA)	0.63	0.15 to 1.61
HOURS (log of actual audit hours)	298	38 to 1,610
YEAR1 (1 = first year engagement, 0 otherwise)	(0 = 168, 1 = 13)	0 to 1
PCI (log of per capita income)	6,251	2,642 to 12,827
CAFR (1 = CAFR, 0 otherwise)	(0 = 159, 1 = 19)	0 to 1
ICRREPORT (1 = reported weaknesses in internal control, 0 otherwise)	(0 = 100, 1 = 78)	0 to 1
TENURE (tenure of the auditor)	10.28	1 to 40
QCR Index** (sum of deficiencies found in QCR)	-37	-131 to -1
DESK Index (sum of deficiencies found in desk review)	-5	-17 to 0
QCRSCORE (based on QCR Index)	0.76	0.01 to 0.99
DESKSCORE (based on Desk Index)	0.72	0.06 to 1.00

*Means and ranges are for true values, not log values.

**QCR Index is the metric used in the prior TEA studies (e.g., Deis and Giroux 1992 & 1996).

Table 5: Pearson Correlation Coefficients

	DEKSCORE	QCRSCORE	SIZE	WEALTH	CLIENTS	OFFICES	PEER	TIME	HOURS	YEAR1
DEKSCORE	1.00									
QCRSCORE	0.16	1.00								
SIZE	0.15	0.07	1.00							
WEALTH	0.02	-0.23	-0.32	1.00						
CLIENTS	0.20	0.19	0.20	-0.11	1.00					
OFFICES	0.11	0.20	0.10	-0.10	0.30	1.00				
PEER	-0.10	0.23	0.19	-0.14	0.09	0.01	1.00			
TIME	-0.14	-0.19	-0.03	0.00	-0.26	-0.04	-0.08	1.00		
HOURS	0.02	0.27	0.63	-0.22	0.08	0.16	0.15	0.10	1.00	
YEAR1	-0.21	0.14	-0.04	-0.04	0.00	-0.08	-0.06	0.12	0.05	1.00

Table 6 Audit Quality Model Based on Desk Reviews
(Dependent variable: DESKSCORE)

Variable	Predicted Sign ¹	Regression Coefficient	T-statistic ²
INTERCEPT		0.424	1.356
SIZE	-	0.024	1.701
WEALTH	-	0.021	0.962
CLIENTS	+	0.008	1.734**
OFFICES	+	0.009	0.640
PEER	+	0.033	0.766
TIME	-	-0.001	-0.833
HOURS	+	-0.024	-0.876
YEAR1	+	-0.142	-2.417***
R-square = .110			
F-value = 2.600**			

¹Predictions according to Deis and Giroux (1996, 72) model
²Significant at the *.10 level; **.05 level; ***.01 level

Deis and Giroux (1996, 69) developed a model of audit hours to investigate the determinants of auditor effort. They found that QCRSCORE was related to audit hours, and suggested that the differences in audit quality measured by QCR scores measured product differentiation in the audit market. That is, for audits to be of high quality, auditors had to expend more effort. We replicate their model adding DESKSCORE (along with QCRSCORE) to measure product differentiation (Table 8). In this study, DESKSCORE is not a significant predictor of audit hours. Again, desk reviews do not seem to capture the differences in audit quality that are measured by QCRs.

Because there is a potential simultaneous equation bias among the DESKSCORE, QCRSCORE, and HOURS models, the three equations were jointly estimated using the three-stage least squares approach. To keep the three-equation system from being over-identified, the equations vary slightly from the OLS models. The three-equation system that was jointly estimated is as follows:

$$DESKSCORE = \alpha_{10} + \beta_{11}SIZE + \beta_{12}WEALTH + \beta_{13}CLIENTS + \beta_{14}OFFICES + \beta_{15}CAFR + \beta_{16}PEER + \beta_{17}TIME + \beta_{18}HOURS + \beta_{19}YEAR1 + \gamma_{11}HOURS + \gamma_{12}QCRSCORE + \epsilon_{11}$$

**Table 7 Audit Quality Model Based On Working Paper Reviews
(Dependent variable: QCRSCORE)**

Variable	Predicted Sign ¹	Regression Coefficient	T-statistic ²
INTERCEPT		1.099	3.789***
SIZE	-	-0.046	-3.483***
WEALTH	-	-0.059	-2.922***
CLIENTS	+	0.004	0.912
OFFICES	+	0.024	1.731**
PEER	+	0.105	2.643***
TIME	-	-0.002	-2.842***
HOURS	+	0.109	4.365***
YEAR1	+	0.142	2.581***
DESKSCORE		0.170	2.391***

adj. R-square = .291

F-value = 7.648***

¹Predictions other than DESKSCORE according to Deis and Giroux (1996, 72) model

²Significant at the *.10 level; **.05 level; ***.01 level

Table 8 Audit Hours Model (Dependent variable: Log of Actual Audit Hours)

Variable	Predicted Sign ¹	Regression Coefficient	T-statistic ²
INTERCEPT		4.709	3.078***
SIZE	+	0.314	10.402***
PCI	-	-0.240	-1.405*
OFFICES	+	0.072	1.821**
CLIENTS	-	-0.022	-1.845**
CAFR	+	0.318	2.300**
ICREPORT	+	0.128	1.462*
OPIN	+	0.236	1.269
YEAR1	+	0.115	0.711
DESKSCORE	-	-0.214	-1.011
QCR	+	0.744	3.673***

R-square = .511

F-value = 17.455***

¹Predictions other than for DESKSCORE according to Deis and Giroux (1996, 69) model.

²Significant at the *.10 level; **.05 level; ***.01 level

$$QCRSCORE = \alpha_{20} + \beta_{21}SIZE + \beta_{22}WEALTH + \beta_{23}CLIENTS + \beta_{24}OFFICES + \beta_{25}PEER + \beta_{26}TIME + \beta_{27}HOURS + \beta_{28}YEAR1 + \gamma_{21}HOURS + \gamma_{22}DESKSCORE + \epsilon_{21}$$

$$HOURS = \alpha_{30} + \beta_{31}SIZE + \beta_{32}PCI + \beta_{33}CAFR + \beta_{34}ICREPORTS + \beta_{35}OPIN + \beta_{36}YEAR1 + \gamma_{31}DESKSCORE + \gamma_{32}QCRSCORE + \epsilon_{31}$$

Table 9 Three-Stage Least Squares Simultaneous Equations Model of Audit Service Supply

<i>Dependent Variable:</i> Variable	DESKSCORE		QCRSCORE		HOURS	
	Regression Coefficient	T-statistic ¹	Regression Coefficient	T-statistic ¹	Regression Coefficient	T-statistic ¹
INTERCEPT	1.202	5.043***	-2.953	-5.617***	4.122	3.966***
SIZE	0.117	4.142***	-0.286	-4.386***	0.354	9.021***
WEALTH	0.007	0.399	-0.014	-0.382	na	na
PCI	na	na	na	na	-0.063	-0.426
CLIENTS	0.002	0.574	-0.006	-0.585	na	na
OFFICES	0.001	0.052	-0.001	-0.042	na	na
CAFR	-0.001	-0.115	na	na	0.071	0.744
ICREPORT	na	na	na	na	0.052	0.535
OPIN	na	na	na	na	0.045	0.426
PEER	0.008	0.277	-0.019	-0.314	na	na
TIME	-0.000	-0.117	0.000	0.162	na	na
YEAR1	-0.136	-2.070**	0.329	2.053**	-0.271	-1.002
<i>Endogenous Variables:</i>						
HOURS	-0.315	-4.181***	0.769	4.572***	na	na
QCRSCORE	0.412	9.970***	na	na	1.155	4.258***
DESKSCORE	na	na	2.429	8.958***	-2.194	-1.749**

¹Significant at the *.10 level; **.05 level; ***.01 level
na: not applicable

The results are generally similar to the OLS results except that all three propositions were upheld with slightly stronger results. In the two audit quality models the endogenous variables DESKSCORE and QCRSCORE are significant and positively associated with one another. Audit hours, however, is positively associated with QCRSCORE but negatively associated with DESKSCORE. This is the same as the OLS result. In the audit hours model QCRSCORE is significant and positively associated with HOURS while DESKSCORE is significant and negatively associated with HOURS. This is likely driven by the “window dressing” phenomena noted in the discussion of Figure 1.

Conclusions

This study investigated ex-post measurements of audit quality in the public sector. The results suggest that desk reviews differ from QCRs in significant ways. Two variables that significantly affect desk review score are the number of ISD clients and the year of the audit (i.e., initial versus continuing audit engagements). The findings indicate that initial audits produce lower desk review scores (more deficiencies), and auditors with more ISD clients receive higher desk review scores (fewer deficiencies). These results can be explained by a new interpretation of desk review scores. Desk reviews may capture the industry/client specific knowledge of the auditor. Finally, because the number of ISD clients and DESKSCORE are not highly correlated, it appears that these two variables measure different forms of industry/client specific knowledge.

The study also finds that desk review score is not a significant predictor of audit hours. Prior research indicated that measures of audit quality based on desk reviews were associated with auditor labor (O’Keefe et al. 1994). Our paper is the first research, however, to use a direct measure of labor (i.e., audit hours) to test this theory. Further, this is the first study to directly compare desk review scores with established and accepted measures of public-sector audit quality. We find that desk review scores are not related to audit hours and do not measure the audit quality characteristics captured by quality control reviews.

It is critical for regulators to learn if QCRs are redundant measures of audit quality, or if they provide valuable information not revealed by desk reviews. Many audits in the public sector are of low quality and fail to comply with GAAS (US GAO 1986). As a result, some type of formal review process is necessary to promote quality financial reporting by governmental units. The TEA and other school districts could significantly reduce public spending if performing only desk reviews would accomplish the monitoring goals set by the Department of Education. This research indicates that desk reviews only capture one aspect of the quality of audit services. Further, study of the indicators of audit service quality in the public sector begins to speak to recent problems with consumer confidence in private sector audit quality. Detailed reviews of working paper evidence can unveil substandard audits that use “window dressing” techniques to conceal their deficiencies.

Limitations and Future Research

The study has limitations related to the database used. QCR results, documented in the three Texas school district studies, were based on a database of audits conducted in 1986 and 1987. Given the comparison goals of this study, the same database was used with the addition of the desk review measure. While we believe that this data can still improve our understanding of audit quality constructs, we must acknowledge that using older data has some limitations. Also, it is possible that the desk review process in Texas differs from the process used in other states. Before results can be generalized to a nationwide population of school districts, QCRs should be compared to desk reviews for school districts of other states. A further limitation is the scoring method used to develop the DESKSCORE metric. Although the scoring method was rigorous and based on TEA guidelines, any weighted scoring procedure is subjective in nature.

As the results indicate, desk reviews and QCRs measure very different constructs. We must investigate whether or not it is beneficial for regulators to provide desk reviews for all school districts and QCRs only on a limited basis. Future research could focus on the particular aspects of auditor expertise captured by the desk review process. In addition, if desk review procedures simply indicate industry specific knowledge of the auditor, we must determine if the desk review process is beneficial. If desk reviews can be replaced by checking specific auditor qualifications, will Texas need to implement QCR programs for all school districts?

References

1. Colbert, G., and T. O’Keefe. 1995. Compliance with GAAS reporting standards: Evidence from a positive enforcement program. *Auditing: A Journal of Practice & Theory* 14 (Fall): 1-16.
2. Copley, P., M. Doucet, and K.M. Gaver. 1994. A simultaneous equations analysis of quality control review outcomes and engagement fees for audits of recipients of federal financial assistance. *The Accounting Review* 69 (January): 244-256.
3. DeAngelo, L. 1981a. Auditor independence, ‘Low Balling,’ and disclosure regulation. *Journal of Accounting and Economics* 3 (2): 113-127.
4. _____. 1981b. Auditor size and quality. *Journal of Accounting and Economics* 3 (4): 183-199.
5. Deis, D., and G. Giroux. 1992. Determinants of audit quality in the public sector. *The Accounting Review* 67(3): 462-479.
6. _____. 1996. The effect of auditor changes on audit fees, audit hours, and audit quality. *Journal of Accounting and Public Policy* 15(1): 55-76.
7. _____, and T. Canby. 1990. Auditing the auditors. *Today’s CPA* 16(3): 36-39.
8. Francis, J., and D. Simon. 1987. A test of audit pricing in the small-client segment of the U.S. audit market. *The Accounting Review* 62(1): 145-57.
9. Giroux, G., D. Deis, and B. Bryan. 1995. The effect of peer review on audit economics. *Research in Accounting Regulation* 9: 63-82.
10. Nichols, D., and K. Price. 1976. The auditor-firm conflict: An analysis using concepts of exchange theory. *The Accounting Review* 51(2): 335-46.
11. O’Keefe, T., R. King, and K. Gaver. 1994. Audit fees, industry specialization, and compliance with GAAS reporting standards. *Auditing: A Journal of Practice and Theory* 13(2): 41-55.

12. O’Keefe, T., and P. Westort. 1992. Conformance to GAAS reporting standards in municipal audits and the economics of auditing. *Research in Accounting Regulation* 6 (1992): 39-77.
13. Pachelles, M., and E. MacDonald. 1999. Ernst & Young woes continue despite Cendant settlement. *The Wall Street Journal Interactive Edition* (December 20).
14. Rubin, M. 1988. Municipal audit fee determinants. *The Accounting Review* 63(2): 219-236.
15. Shockley, R. 1982. Perceptions of audit independence: A conceptual model. *Journal of Accounting, Auditing, and Finance* 5: 126-143.
16. Simon, D., and J. Francis. 1988. The effects of auditor change on audit fees: Tests of price cutting and price recovery. *The Accounting Review* 63 (2): 255-269.
17. Texas Education Agency. 1994a. *Audit Review Manual*. Austin, TX.
18. _____. 1994b. *Checklist for Internal Review*. Austin, TX.
19. Turpen, R. 1990. Differential pricing on auditor’s initial engagements: Further evidence. *Auditing: A Journal of Practice and Theory* 9 (Spring): 60-76.
20. United States General Accounting Office. 1986. *CPA Audit Quality: Governmental Audits Do Not Comply with Professional Standards*. Washington, D.C.: GAO.

Endnotes

¹ One of the objectives of both desk reviews and working paper reviews is to decide whether the auditor’s work is acceptable or unacceptable (i.e., substandard). This is commonly referred to as the “outcome” of the review process. Several studies have used the outcome of the review, rather than the underlying detail, to proxy for audit quality. Colbert and O’Keefe (1995) and O’Keefe and Westort (1992) use the outcome of desk reviews conducted by the Oregon State Board of Public Accountancy. Copley et al. (1994) use the outcome of working paper reviews conducted by Offices of Regional Inspector’s General. A dichotomous dependent variable is used in each of these studies to indicate acceptable versus unacceptable outcomes.

² Deis and Giroux (1992) also assigned different weights to major and minor deficiencies in quality control reviews. Results of the current study are robust with multiple weight assignments.

³ Diagnostic measures were performed for all models. Variance inflation factors were all below 2, indicating few problems with multicollinearity. Residual plots indicated that residuals were normally distributed for all models. Plots of residuals versus predicted values were analyzed to check for constant variances. All models had homoskedastic errors after applying log transformations to some variables. Whites Specification Test also indicated that the residual terms were homoscedastic.

⁴ To verify that the non-significant results for Audit hours were not the result of differences between our limited sample of 181 school districts and the full sample of 232 school districts investigated by Deis and Giroux (1996), we re-estimated the Deis and Giroux model for audit quality using our limited sample. Results were unchanged from their original study. Audit hours had a positive and significant effect on audit quality as measured by QCRSCORE.