RESEARCH IN GOVERNMENTAL AND NONPROFIT
ACCOUNTING VOLUME 11

RESEARCH IN GOVERNMENTAL AND NONPROFIT ACCOUNTING

EDITED BY

PAUL A. COPLEY

School of Accounting, James Madison University, USA

2004

ELSEVIER
JAI

Amsterdam – Boston – Heidelberg – London – New York – Oxford
Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo

RESEARCH IN GOVERNMENTAL AND NONPROFIT ACCOUNTING

RESEARCH IN GOVERNMENTAL AND NONPROFIT ACCOUNTING

Series Editor: Paul A. Copley

Volumes 1–9: Edited by James L. Chan
Volume 10: Edited by Paul A. Copley

EDITORIAL BOARD

ELSEVIER B.V.
Sara Burgerhartstraat 25
P.O. Box 211
1000 AE Amsterdam
The Netherlands

ELSEVIER Inc.
525 B Street, Suite 1900
San Diego
CA 92101-4495
USA

ELSEVIER Ltd
The Boulevard, Langford
Lane, Kidlington
Oxford OX5 1GB
UK

ELSEVIER Ltd
84 Theobalds Road
London
WC1X 8RR
UK

First edition 2004

Library of Congress Cataloging in Publication Data
A catalog record is available from the Library of Congress.

British Library Cataloguing in Publication Data
A catalogue record is available from the British Library.

ISBN: 0-7623-1123-1
ISSN: 0884-0741 (Series)

♾ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper). Printed in The Netherlands.

CONTENTS

LIST OF CONTRIBUTORS

Arthur Allen	University of Nebraska-Lincoln
Earl D. Benson	Western Washington University
Edmund J. Boyle	University of Rhode Island
Steven M. Cooper	Bureau of Audits, State of Rhode Island
Edward B. Douthett, Jr.	George Mason University
Randal J. Elder	Syracuse University
Mary L. Fischer	University of Texas, Tyler
Marshall A. Geiger	University of Richmond
Gary Giroux	Texas A&M University
Janet S. Greenlee	University of Dayton
David Gotlob	Indiana University – Purdue University (deceased)
Sarah A. Holmes	Texas A&M University
Laurence E. Johnson	Colorado State University
Rowan Jones	University of Birmingham, UK
Carol M. Lawrence	University of Richmond
Suzanne H. Lowensohn	Colorado State University
Barry R. Marks	University of Houston, Clear Lake
Nicholas P. Marudas	Auburn University, Montgomery
Andrew J. McLelland	Auburn University
Carolyn Strand Norman	Virginia Commonwealth University
Linda M. Parsons	George Mason University

EDITOR'S COMMENTS

I am pleased to present Volume 11 of *Research in Governmental and Nonprofit Accounting*, the only series dedicated exclusively to governmental and nonprofit accounting and reporting issues. The purpose of *Research in Governmental and Nonprofit Accounting* is to stimulate and report high-quality research on a wide range of governmental and nonprofit accounting topics.

The first three articles deal with issues related to the bond market. Individuals seeking to familiarize themselves with this literature will find the Reck, Wilson, Gotlob and Lawrence article particularly useful. The *Benson and Marks and Allen and Sanders* articles provide insights for those seeking to design bond studies.

The private nonprofit sector continues to be an area of high research activity as data becomes increasingly available. Similarly, international issues are an emerging area of research and the *Giroux, McLelland and Jones* article presents a comparative analysis between U.S. and U.K. governments.

The issue of audit quality has become particularly important in recent years with five related articles appearing in Volume 11. The 2003 revision of the Yellow Book and the independence standards in particular should provide opportunities for continued research in the area of public sector auditing. *Marc Rubin* observes a lack of research in public school performance measurement and reporting. I am also optimistic that this sector will provide many research opportunities in the future. The *Patten and Strong* article is an early response to this call for research.

GOVERNMENTAL CAPITAL MARKETS RESEARCH IN ACCOUNTING: A REVIEW, EXTENSION, AND DIRECTIONS FOR FUTURE RESEARCH

Jacqueline L. Reck, Earl R. Wilson, David Gotlob (deceased) and Carol M. Lawrence

ABSTRACT

We present a review of the governmental capital markets research that has been conducted since Ingram et al. (1987). The review indicates that much of the research conducted in the last fifteen years involves the new issue bond market. Increased attention is also given to research in the secondary market. However, much of the secondary market research fails to examine the specific relationship between secondary market returns and accounting information. To that end, we provide evidence of the relationship between secondary market bond yields and two accounting variables used in prior studies. Our results indicate that accounting information is significantly associated with the yields of infrequent bond issuers, supporting research conducted in the new issues market.

Given the many policy developments related to financial reporting and disclosures that have occurred over the past 10–15 years, additional

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 1–33

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11001-X

research is needed in the government capital markets area to examine whether these policies have improved the usefulness of governmental accounting information. Therefore, throughout the paper, we provide direction for future government capital markets research.

1. INTRODUCTION

It has been over a decade since Ingram et al. (1987) provided a review of the capital markets research in government accounting. Since that time additional research in the governmental capital markets area has been conducted, and significant policy developments have occurred that raise new financial reporting and disclosure issues and provide unique opportunities for additional research in this area. Consider, for example, that institutional changes in governmental accounting and auditing have dramatically improved the quality of governmental accounting and auditing since the main body of this research was published.[1,2] Thus, an important empirical question is whether improvement in the quality of governmental financial reporting has strengthened the association between municipal bond measures and accounting and reporting information.

More recently, the Securities and Exchange Commission (SEC), through its Municipal Securities Rule-Making Board (MSRB), significantly increased debt issuers' disclosure requirements to municipal bond investors.[3] As a result, there is now a strong need for additional research regarding the decision usefulness of recent regulatory policies aimed at overcoming deficient financial reporting quality, and continuing disclosure to investors in the municipal bond market.

Finally, the Governmental Accounting Standards Board's (GASB) new governmental reporting model (GASB, 1999) represents a revolutionary change in the history of governmental accounting. The new model, when fully implemented, will provide many additional research issues, such as whether government-wide financial information similar to that presented by for-profit organizations is more useful in pricing bonds than the traditional fund-based financial information currently reported, and whether disclosure of infrastructure assets is useful in pricing debt.

The purposes of this paper are three-fold: (1) to review the research over the past 15 years that investigates the impact of accounting related information on the municipal bond market; (2) to provide additional empirical research related to one of the more thinly studied areas – the municipal bond secondary market; and (3) to provide direction for future research.

In the next section we review the research published since Ingram et al. (1987) and provide direction for future research. An empirical analysis of the issue of whether government-specific financial accounting information is impounded in secondary market prices is presented in Section 3. Concluding remarks are presented in Section 4.

2. A REVIEW OF RECENT GOVERNMENT CAPITAL MARKETS RESEARCH

This review extends Ingram et al. (1987), which focuses on the association between accounting and audit information and bond market measures. The bond measures in their study include bond ratings, new issue costs and seasoned market yields. In this section we review the subsequent governmental capital markets research, dividing the studies according to the same bond measures – bond ratings, new issue costs and seasoned market yields. These three areas are further divided based on the effects of financial information and reporting practices, specific accounting issues, and audit reports/practices. Table 1 provides a summary of the additional research that has been conducted in each of these areas.

Table 1. Government Capital Markets Research in Accounting since 1985.

	Bond Ratings	New Issues	Seasoned Issues
Financial information/reporting practices	Lewis, Patton and Green (1988) Wescott (1988)	Wilson and Stewart (1990) Benson, Marks and Raman (1991) Fairchild and Koch (1998) Stover (1991) Feroz and Wilson (1992) Soybel (1992) Vijayakumar (1995)	Ingram, Raman and Wilson (1989) Liu and Seyyed (1991) Marquette and Wilson (1992) Ingram and Wilson (1999)
Specific accounting issues		Copeland and Wilson (1986) Chaney (1998) Sneed and Sneed (1999)	Raman and Wilson (1990)
Audit reports/practices	Langsam and Kreuz (1990)		Raman and Wilson (1994)

2.1. Bond Rating Research

At the time of the Ingram et al. (1987) review article, more studies (e.g. Copeland & Ingram, 1982; Ingram & Copeland, 1982a; Marquette & Marquette, 1986; Raman, 1981, 1982a, b, 1986; Wallace, 1981; Wescott, 1984; Wilson & Howard, 1984) had examined bond ratings than any other municipal bond market measure. These studies find an association between accounting variables and bond ratings and bond rating changes. In general, however, the predictive accuracy for municipal bond ratings models is about 10–20% less than that for corporate bond rating prediction models. The authors (Ingram et al., 1987) suggest that the observed difference is attributable to noise in government accounting data. This argument is plausible since the studies cited by Ingram et al. (1987) use data from the period prior to the formation of the GASB, implementation of the Single Audit Act, and the SEC disclosure requirements.[4] Poor governmental accounting and reporting practices were well documented during the late 1970s and early 1980s. Additionally, many of the studies reviewed by Ingram et al. (1987) introduce measurement error since they rely on accounting data obtained from the Bureau of Census rather than from actual annual reports. Icerman and Welch (1989) find that there are differences of approximately 7% between Census-based financial information and information collected directly from the municipality's annual report.

As can be seen from Table 1, relatively little accounting research has focused on bond ratings since Ingram et al. (1987). One reason is the increasing availability of bond price data over the past 10–15 years. Since prices are continuous measures of default risk and marketability they are preferable for research purposes over the crude ordinal bond rating measure. Additionally, the increasing availability of bond price data has allowed researchers to investigate bond market participants other than bond rating agencies, such as underwriters, secondary market investors, and bond insurers.

2.1.1. Financial Information/Reporting Practices

Because prior research (e.g. Marquette & Marquette, 1986) finds that environmental variables may impact the usefulness of financial accounting information in assessing bond ratings, Wescott (1988) uses cluster analysis to divide her sample into three groups of cities with similar environmental characteristics. For each of the three city groups defined in her study, Westcott conducts a probit analysis using financial accounting variables to predict bond ratings. She finds that the significance of financial accounting variables varies by city group, implying that for different market segments accounting information may be impounded differently. Classification accuracy for Wescott's (1988) models is greater than 70% for two of the three city groups. However, classification accuracy ranges

only from 28 to 53% when the probit model is used to predict ratings of a holdout sample. The large number of variables, combined with small sample size may contribute to Westcott's inability to construct a generalizable predictive model.

Recently, only one other accounting study has attempted to look at bond rating behavior. Lewis et al. (1988) use a judgment-based lens model to compare financial analysts' information choices for rating a bond with predictions of a statistical bond rating model. They provide analysts with twelve items of information, nine of which are financial accounting variables and three of which are environmental variables. Analysts are allowed to select five items[5] on which to base their rating decision. The results indicate that analysts' predictive accuracy (39%) is comparable to the statistical models (41%) when original selections and weights are applied across time periods (Lewis et al., 1988). A limitation of their study is that variable choices are imposed on the analysts. Additionally, since there is no theory on which to base accounting variable selection, the nine accounting variables provided to subjects may not be the most salient variables. Interestingly, the statistical models do not select any of the environmental variables, but rely solely on financial accounting variables (Lewis et al., 1988).

Both of the above studies (Lewis et al., 1988; Wescott, 1988) rely on data from the early 1970s to the early 1980s. During this time financial information of government municipalities was less standardized and less reliable than the financial accounting and reporting data that has become available since the inception of GASB and the Single Audit Act. Since the financial data may have been viewed as unreliable by users, the ability to construct predictive models was seriously impacted. Additionally, both studies rely on Census-based financial data, which has been shown (Icerman & Welch, 1989) to vary from data reported in actual financial statements.

2.1.2. Audit Reports/Practices

Langsam and Kreuz (1990) look specifically at the impact of audit opinion on bond measures. They find a significant association between audit opinion and bond ratings. Bond rating is a dichotomous variable, represented as high quality bond ratings (Aa and above) and lower quality ratings. The study relies on a simple correlation between audit opinion and bond rating. Because bond ratings are divided into two groups, the association between opinion and each level of bond rating cannot be determined. Furthermore, it is unknown whether the association holds if financial accounting control variables are included in the analysis.

2.1.3. Summary and Areas for Future Bond Rating Research

In general, the few bond ratings studies conducted within the past 10–15 years reinforce the findings of the earlier research. The more recent studies find that

accounting and auditing information are significantly associated with bond ratings. The main contributions of the more recent research are the introduction of new methodologies (cluster analysis and behavioral research) and examination of audit variables, such as audit opinions, in addition to accounting variables.

A recent study by Fitch IBCA (Bond Buyer Staff, 2000) indicates that municipal bonds may be systematically underrated. As a result, Fitch IBCA reassessed its municipal bond rating process. Additionally, Moody's Investors Service indicated it will work to standardize its rating process (Bond Buyer Staff, 2000). The changes being implemented by the bond rating agencies may have implications for what and how financial and non-financial information is incorporated into the bond rating process. For example, Fitch IBCA indicated that increased emphasis will be placed on management practices (Bond Buyer Staff, 2000). Additionally, a more objective emphasis will be placed on debt and financial performance. GASB Statement No. 34 (GASB, 1999) requires, for the first time, that government annual reports include a Management Discussion and Analysis (MD&A). We expect that rating analysts will find the MD&A, which provides information useful in assessing management performance, (and other features of the new model such as government-wide financial information prepared on the accrual basis) useful in evaluating the creditworthiness of municipal bonds. As the reliability of accounting information continues to improve, the predictive ability of bond ratings models should improve as well.

2.2. New Issues (Primary) Bond Market

Since Ingram et al. (1987), a fair amount of government capital markets research has been conducted in the new issues, or primary bond market. The research includes a variety of financial accounting variables and reporting practices, as well as new accounting issues. No recent research has been published involving the effects of audit reports and practices on new issue yields.

2.2.1. Financial Information/Reporting Practices

Ingram et al. (1987) reviewed prior accounting research that examines the importance of accounting/auditing information and the quality of financial reporting in pricing new bond issues. Those studies (Apostolou et al., 1985; Wallace, 1981; Wilson, 1983; Wilson & Howard, 1984) produce mixed results concerning the association between accounting variables and municipal bond yields and net interest costs or yields. Not addressed by Ingram et al. (1987) are two studies (Benson et al., 1984, 1986) that find that variables reflecting financial reporting quality (specifically, state imposed accounting requirements and the Government

Finance Officers Association's Certificate of Achievement) are significant in explaining yields and returns under certain conditions. In the time period since 1987, researchers continue to examine the association between quality of reporting and new issue measures. Researchers have also identified additional factors that may impact the use of accounting information – perhaps helping to explain the inconsistent results in the studies cited by Ingram et al. (1987).

If financial accounting information is not considered reliable or relevant, uncertainty arises concerning the quality of management and the government's financial situation. This may lead to higher (net) interest costs on new issues. The question of relevant and reliable information is addressed in two research studies (Benson et al., 1991; Fairchild & Koch, 1998). The Benson et al. (1991) study investigates whether voluntary disclosures by states impact the state's net interest cost. Their results indicate that the net interest cost for high disclosure states is on average 14 basis points lower than the net interest costs for low disclosure states. These results are supported by Fairchild and Koch (1998), who look at the relationship between municipal disclosure practices and net interest costs. They find that for unrated bonds there is a differential of approximately 14 basis points between municipalities in states requiring high disclosure in official statements vs. municipalities in states with low disclosure requirements for official statements. These results are surprisingly similar given the fact that Benson et al. (1991) rely on data from October 1976 to May 1979 while the Fairchild and Koch (1998) paper utilizes data from 1980 to 1988. Furthermore, the method used to identify high disclosure vs. low disclosure observations is somewhat different. The first study (Benson et al., 1991) relies on a six-point index based on twelve accounting practices; whereas, the second study (Fairchild & Koch, 1998) relies on a dichotomous variable based on a late 1980s survey conducted by the National Association of State Auditors, Comptrollers, and Treasurers.

Both the Benson et al. (1991) study and the Fairchild and Koch (1998) study rely on a considerable amount of data from the pre-GASB, pre-Single Audit, and pre-SEC disclosure period. As indicated earlier, accounting information and reporting practices have greatly improved since the mid-1980s. While the question of voluntary disclosure is still highly relevant, state regulation of disclosure may have less impact on new issue costs as accounting and reporting practices become increasingly more standardized nationwide. Additionally, the findings of studies in this area may be confounded if high disclosure states are found to have a higher proportion of frequent issuers, since lower average information asymmetry, and thus, lower borrowing costs would exist in these states, even without reporting regulation.

Closely related to the quality of disclosure studies is a study conducted by Soybel (1992) in which she assesses the association between new issue yields

and cash basis accounting vs. GAAP (modified accrual) accounting in the case of New York City. She finds that the market is not deceived by the use of cash basis revenues and expenditures; rather the market appears to adjust to GAAP based accounting numbers for setting new issue yields. The results of Soybel's (1992) study indicate market efficiency, in that the market is not deceived by New York City's overstatement of revenues and understatement of expenditures. The generalizability of Soybel's study is somewhat limited given the city studied. Because New York City is the largest city in the United States, it has characteristics that may not be relevant to other U.S. cities. Also, since New York City was reporting on a cash basis, Soybel was required to convert to the GAAP basis. The conversion process was necessarily crude and contains measurement error since not all financial accounting records were available to Soybel.

Wilson and Stewart (1990) and Vijayakumar (1995) show that the level of information asymmetry varies across segments of the new issue bond market. Wilson and Stewart (1990) are the first to examine the effects of information asymmetry on the association between financial disclosure and competition among underwriters for new bond issues. They find that financial accounting disclosure is more important to the market competitiveness (number of bids received on the debt offering) of infrequent debt issuers than to the competitiveness of frequent debt issuers. Vijayakumar (1995) also finds that information asymmetry may cause "high quality" bond issuers to include a call provision, since the issuer is unable to credibly reveal its financial position to the markets.

Feroz and Wilson (1992) argue that the primary bond market could be divided into a high information asymmetry market segment consisting of the bond issues of smaller, less frequent issuers that are managed by local or regional underwriters, and a low information asymmetry segment managed by national underwriters. They then hypothesize that financial accounting and auditing information is more value-relevant to the market involved with new issues underwritten by local or regional underwriters than issues underwritten by national underwriters. They find that financial accounting disclosures, type of auditor, and quality of financial management are all significantly associated with net interest costs if the bond issue is managed by a local or regional underwriter. None of these variables are significant if the issue is managed by a national underwriter.

The fact that size matters in determining net interest costs is further supported by Reeve and Herring (1986), who find that there is an average difference of 40 basis points between the net interest costs of small unrated bond issuers and large unrated bond issuers. Costs this large indicate that additional information disclosure is beneficial for smaller issuers.

The effect of intermediation, discussed in more detail in the seasoned bond market, is also touched on in the new issues market. Stover (1991) uses path

analysis to investigate the relationship between bond ratings and debt related financial variables and new issue reoffering yields. He finds that debt variables have no direct effect on reoffering yields, but instead exert an indirect effect through the entity's bond rating. Again, Stover (1991) relies on 1980s data in his tests. As the quality and amount of accounting data has improved, it is possible that users are more willing to evaluate and incorporate information directly into the yields, rather than relying on evaluation by bond raters.

2.2.2. Specific Accounting Issues

Only two studies examined by Ingram et al. (1987) look at the impact of a specific accounting issue on the yields or returns of new bond issues. Both studies (Copeland & Ingram, 1983; Marks & Raman, 1985) investigated the impact of pension ratios on net interest costs. The results of the studies are mixed, possibly due to the poor quality of the information concerning unfunded pension obligations. At the time, the studies relied either on pension ratios or pension funding estimates made by Kotlikoff and Smith (1983), since government entities were not required to report unfunded pension obligations.

Copeland and Wilson (1986) use pension data reported by 92 cities in determining if pension contributions influenced new issue yield premiums. The results of the study indicate that pension funding practices influence new issue yield premiums.

A later study by Sneed and Sneed (1999) improves upon earlier pension research by using actual unfunded pension obligation data reported by states in their annual report footnotes. The collection of unfunded pension obligation information was made possible through implementation of GASB Statement No. 5, *Disclosure of Pension Information by Public Employee Retirement Systems and State and Local Government Employers* (GASB, 1986). Implementation of GASB Statement No. 5 improved upon the reporting of pension costs used by Copeland and Wilson (1986) since it allowed for increased consistency and comparability of reported pension information. The results of Sneed and Sneed's (1999) study indicate that unfunded pension obligations have a significant positive association with new issue yields to maturity for bonds with maturities greater than ten years. For relatively short terms to maturity (ten years or less), the unfunded pension obligation is not significant. Additional research is needed to examine whether Sneed and Sneed's results hold with the changes required in pension accounting by GASB Statement Nos. 25, 26 and 27 (GASB, 1994a, b, c).

In addition to the research on pensions, Chaney and Copley (2001) investigate the impact of GASB Statement No. 14 (GASB, 1991) on new issue costs. Specifically, Chaney and Copley look at whether reporting the financial and/or control relationships between the primary government and its component units is impounded in true interest costs of a new issue. Chaney and Copley (2001)

hypothesize that the reporting of financial relationships is important to the market since it helps the market determine the entity's default risk. While results concerning the importance of financial relationships between the primary entity and component units are inconclusive, Chaney and Copley (2001) find support for a stronger association between new issue costs and combined reporting than between new issue costs and other types of reporting; thus, providing support for GASB Statement No. 14.

2.2.3. Summary and Areas for Future Research in the New Issues Market

In general, the recent research using new issues market data provides stronger evidence of the association between accounting related information and new issue measures than is provided by the earlier research in the Ingram et al. (1987) review. In the studies involving specific accounting issues, it appears that changes in financial reporting are associated with lower borrowing costs on new bond issues. Additionally, the more recent research in this area seems to indicate that the relationship between accounting information and new issue variables is complicated by various institutional factors (such as market segmentation and information asymmetry).

Much of the research in the new issues market relies on net interest cost (NIC) as the dependent variable. True interest cost (TIC) is a superior measure of new issue cost since it considers the time value of money, and the theory underlying its calculation is the same as that used to support the yield to maturity measures used by some of the secondary market studies (Public Securities Association, 1987). In the past, NIC was primarily used by the market due to the simplicity of calculation. However, with the advances in technology, TIC is growing in popularity (Public Securities Association, 1987). As a result, future researchers need to control for differences that may arise in new issue costs. An empirical question exists as to whether the association between accounting information and new issue costs changes when TIC, rather than NIC, is used in research models.

Future research should also focus on the association between new issue measures and the disclosures required by GASB's new reporting model, such as those in the MD&A. Additionally, if Soybel is correct and the market is efficient in impounding information, a question can be raised concerning the importance of the new GASB reporting model (GASB, 1999) to the government capital market. For example, future research may address whether the relatively greater transparency afforded by the new reporting model improves the efficiency of the market when compared to the "old" reporting model.

Finally, no recent research has examined the effect of audit quality on the new issue market. Changes have occurred with regard to audit and audit quality as a

result of the Single Audit Act of 1984. Future research may address what, if any, impact the changes have on the new issue market.

2.3. Seasoned Issues (Secondary) Bond Market

Ingram et al. (1987) cite five studies related to the secondary municipal bond market. Only one of the studies (Ingram, 1983) looks at the significance of accounting variables on secondary market yields. One study (Ingram & Copeland, 1982b) investigates the significance of a regulatory variable (state imposed accounting requirements) on secondary market yields. Two studies (Ingram, 1986; Ingram & Copeland, 1986) investigate the impact of the quality of financial reporting (conformity to the fund accounting model) on municipal bond *risk*. The final study (Ingram et al., 1983) looks at whether a bond rating change provides new information to the secondary market. As can be seen, few studies have specifically addressed the association between accounting information and secondary market yields or returns. Given the paucity of research in this area, and the generally low reliability of the secondary market data used in earlier studies, there continues to exist a need for research regarding the effects of disclosure practices and the usefulness of accounting information in pricing bonds in the municipal bond secondary market. The generally low reliability of secondary market data in earlier research relates to the fact that most of the studies rely on the Standard and Poor's *Blue List*. Ingram et al. (1987), based on empirical evidence from Ingram (1985), suggest that secondary market data from a proprietary pricing service is preferable to secondary market data from the Standard and Poor's *Blue List*.

Since Ingram et al. (1987), six additional accounting related research studies have been conducted in the seasoned issues (secondary) municipal bond market. Thus, compared with the numerous accounting studies published involving corporate securities secondary markets, accounting research in the municipal bond secondary market remains a thinly studied area.

2.3.1. Financial Information/Reporting Practices

Research since Ingram et al. (1987) examines whether seasoned bond prices or yields in the secondary market efficiently impound information. The issue of when seasoned bond yields reflect new financial accounting and reporting information has been studied around three information release times – bond rating changes, release of the annual report, and issuance of new bonds by the same government.

Marquette and Wilson (1992) find that secondary market yields impound information somewhat efficiently. Efficiency is indicated by the fact that secondary

market bond prices significantly respond to bond rating downgrades up to 24 months prior to the actual bond rating change (Marquette & Wilson, 1992). However, there is no significant market response preceding bond rating upgrades.

Other studies specifically look at the time information is released and the impact the information has on seasoned bond prices or yields. Ingram et al. (1989) investigate whether there are significant residual returns around three alternative information release dates – release of the annual financial report, a change in bond rating, and new debt issuance. They find significant residual returns around the bond rating change event and at the time of a new debt issue, but no significant residual return at the time the annual report is released.

Ingram et al. (1989) indicate that the lack of reaction to the annual report information could be related to the high cost of obtaining and using the information relative to its expected value. That is, because of thin trading in the secondary market and the relative difficulty in obtaining annual report information, the market may believe the cost of the information exceeds its value in pricing bonds. Alternatively, the lack of reaction to the annual report information could be related to misspecification of the event date. Unlike the new issue date and the bond rating change date, which are known dates, Ingram et al. (1989) rely on a survey of government officials to identify the annual report release date. Approximately 60% of the respondents to their survey did not know the mailing date of the annual report, and thus provided an estimate (Ingram et al., 1989). Also, more information leakage may occur prior to the annual report release date than prior to the new issue or bond rating date.

Although Ingram et al. (1989) indicate that the market does react to information at the time of a new issue, their study does not address whether financial accounting measures are specifically part of the information associated with seasoned bond yields. That is, the researchers do not test whether the market is reacting to the release of information in general, or whether it is reacting to the release of accounting-based financial information. Liu and Seyyed (1991) find that three revenue and debt variables, along with bond ratings, are significant in explaining yield premiums. Liu and Seyyed (1991) rely on Standard and Poor's *Blue List* data for secondary market information, rather than the more reliable proprietary bond price data used by Ingram (1985). The sample period used by Liu and Seyyed (1991) is quite constrained, covering only three days, one selected from each year 1981, 1982 and 1983.

Another, more extensive, study by Ingram and Wilson (1999) looks at whether financial accounting information is impounded in different bond measures related to bond ratings, new issue interest costs, and seasoned bond yields. Seasoned bond yields are based on data provided by a proprietary pricing service. Financial position and debt are significant in explaining bond ratings, while financial

position and bond rating are associated with net interest costs on new issues. No association is found between accounting information and seasoned bond yields, after controlling for the information impounded in new issue net interest cost. However, for one of their dependent variables, the cumulative percentage yield change, they find that financial position is significant for *larger* cities, even in the presence of net interest costs. The results seem somewhat incongruent given the Feroz and Wilson (1992) results and the results from the corporate literature, which generally indicate that accounting information is relatively more important for smaller entities. The apparent incongruence with the Feroz and Wilson (1992) study and corporate studies, leads, in part, to the analysis presented in Section 3. The fact that part of the Ingram and Wilson (1999) sample is from the pre-GASB, pre-Single Audit, and pre-SEC disclosure period also is partially addressed in Section 3.

2.3.2. Audit Reports/Practices

In 1987 the General Accounting Office (GAO) identified "four critical attributes" for procuring quality government audits, and strongly encouraged the use of the attributes by government entities (Raman & Wilson, 1994). Responding to this policy recommendation, Raman and Wilson (1994) investigate whether quality audits, as identified by the "four critical attributes," are impounded in seasoned bond yield premiums. Due to interdependence among the four attributes, Raman and Wilson (1994) construct a composite procurement variable. When tested, the procurement variable is significantly associated with yield premiums. The significance of the procurement variable appears to be driven by the portion of the sample that received (then) Big 8 audits. An analysis of the components of the procurement variable indicates that a measure of audit competition, the number of audit bids received, is significant for both (then) Big 8 and non-Big 8 audited government entities. Thus, it appears that improving the quality of the audit procurement process is only partially valued by the seasoned bond market. It should be kept in mind that at the time of the Raman and Wilson study only one year of data could be included that specifically incorporated the GAO suggestions.

2.3.3. Specific Accounting Issues

Raman and Wilson (1990) investigate whether pension obligations are impounded in seasoned bond yields. Prior to this study no one had used regression analysis to look at the importance of pension obligations in explaining yields. The results indicate that unfunded pension obligations are impounded by the seasoned market; however, when compared with general obligation debt it appears that the market is only partially impounding the information (Raman & Wilson, 1990). In

general, however, they find that the market appears to price a dollar of unfunded pension obligation similar to a dollar of bonded debt. In the study, Raman and Wilson use the less reliable pension liability estimates provided by Kotlikoff and Smith (1983).

2.3.4. Summary and Areas for Future Research in the Seasoned Issues Market

Secondary market research is focused primarily on the information content of various disclosures. Generally, the findings indicate that information disclosures are valued by the market; however, the high perceived costs vs. benefits of obtaining information appears to affect the market's incorporation of all available information.

Much of the secondary market accounting research is focused on cross-sectional studies conducted during the time period 1978–1990. Of interest would be whether the market differentially impounds financial accounting information during economic recessions and periods of prosperity. If economic recessions increase the likelihood of default risk, and accounting measures are helpful in assessing default risk, one would expect investors to place increased reliance on financial accounting measures during or just prior to economic recessions.

Policy has been implemented which allows for additional access to government financial reports. As a result of a number of financial problems,[6] the Securities and Exchange Commission (SEC) (1994) amended rule 15c2–12, requiring that those municipalities with publicly held debt file their financial reports with an SEC designated repository. Future studies should consider whether SEC policies have improved the usefulness of accounting data to the market. For example, the studies that address the market segmentation issue rely on information asymmetry, resulting from cost of information, to explain the existence of market segments. However, the SEC's rule 15c2–12 (1994) requires that bond issuers provide annual filings of financial information to Nationally Recognized Municipal Securities Information Repositories. If research finds that segmentation persists in the presence of such publicly available information it would seem to indicate that the cost of obtaining the public information remains too high to eliminate asymmetry.

Finally, additional work remains concerning the information content of accounting variables. The one study (Ingram & Wilson, 1999) that has been conducted focuses on intermediation and the use of accounting information rather than looking specifically at the information content of accounting variables. Due to the lack of research concerning the impact of accounting information on the secondary market and the counterintuitive findings of Ingram and Wilson (see financial information/reporting practices of this section) the following study is presented.

3. AN EMPIRICAL EXAMINATION OF THE ROLE OF GOVERNMENTAL ACCOUNTING INFORMATION IN PRICING SEASONED MUNICIPAL BONDS

As with prior studies (Ingram et al., 1989; Ingram & Wilson, 1999), we find a significant market reaction to entity-specific information released at the time of a new bond issue. However, we go well beyond earlier studies by specifically providing evidence that accounting information is impounded in seasoned bond prices, but only within the market segment of bonds issued by smaller, infrequent issuers, and not the segment of frequent issuers. Our finding is important in that it contradicts the findings of Ingram and Wilson (1999) and provides evidence that Feroz and Wilson's (1992) market segmentation argument also holds in the secondary market for municipal bonds.

The next section provides an institutional perspective on the municipal bond market and discusses the informational efficiency of the primary and secondary bond markets. After providing an institutional perspective, we present the hypotheses. The sample, research design, and model development follow the hypotheses. The final section presents a summary of the study.

3.1. Institutional Background and Prior Research

Given that issuer-specific information is not readily available in the municipal secondary market, particularly for many infrequent issuers, investors face uncertainty about issuers' default risk, and thus, about the precise value of seasoned bonds.[7] In such a market the large amount of information provided in marketing *new* bond issues affords holders of seasoned bonds of the same issuer an opportunity to reassess issuer default risk and revalue seasoned bonds.

Extensive information exchange occurs in the months surrounding the issuance of new bonds in the municipal bond market. Prior to the issuance, the trading desks of all bidding syndicate members contact their regular and potential customers to determine investor interest in a municipality's bonds and possible selling prices (The Bond Market Association, 2001). During the pre-marketing process, analysts in the secondary market can determine whether current price quotes for seasoned bonds are consistent with the price of the new issue, given the widely disseminated new issue information about the municipality.

After the issue has been sold, net interest cost and reoffering yields are published in *The Bond Buyer* and other proprietary sources (The Bond Market Association, 2001). Prices of seasoned bonds can be expected to adjust to the new

issuer-specific information being disseminated to the secondary market, as well as to the new prices themselves.

3.2. Hypotheses

Accounting information is potentially useful for assessing municipal default risk on bonds, and thus, for fine-tuning the price of seasoned bonds. Secondary market investors can be expected to react to the signal conveyed by the *level* of accounting variables at the time of a new issue, given the lack of frequent entity-specific disclosures for most municipal issuers as compared with the regular quarterly disclosures characteristic of corporations.

As suggested previously, the effect on secondary market bond prices of information released during new issues is expected to be stronger for *infrequent* issuers. This expectation is based on theoretical arguments (e.g. Ohlson, 1979) and related empirical evidence (e.g. Atiase, 1985; Feroz & Wilson, 1992; McNichols & Manegold, 1983) that the market reaction to an informational disclosure will be greater, the longer the period since a preceding disclosure.

The inherent complexity and ambiguity of the signal conveyed by municipal accounting information makes it difficult to unambiguously model how bond prices will react to new information. Issuance of new debt does not automatically mean bad news; it may reflect increasing demand for capital assets due to population growth, and revenue capacity may be growing more rapidly than debt service requirements and demand for services. Similarly, although strong current financial position in itself conveys "good news," it may be dominated by contrary economic, social, and political factors, with the result that the overall signal conveyed by the new issue is "bad news." As a result of different values placed on different information, investors may have heterogeneous beliefs about the implications of the overall signal for the issuer's default risk, regardless of the signal conveyed by accounting information.

Such ambiguity in the signal conveyed by municipal disclosures may explain Ingram et al.'s (1989) finding that the *mean* of the bond return distribution is unaffected by new information released in connection with the marketing of new issues, while *variability* increases strongly in response to new information. Nevertheless, we argue that investors generally prefer a signal of strong financial condition to one of weak financial condition, and low debt burden rather than high. Thus, despite the potential for ambiguous signals in municipal disclosure, we expect to observe, on average, a positive association between returns on seasoned bonds and current financial position, after controlling for other factors known to affect bond returns. Similarly, one would expect a negative association

between returns and the relative debt burden. As indicated earlier, we also expect the effect to be influenced by the frequency of issue.

Based on the foregoing arguments, we state the following hypotheses:

H1. The positive association between holding period returns on seasoned bonds around the time of a city's new issue and current financial position of the city is significantly stronger for low frequency issuers than for high frequency issuers.

H2. The negative association between holding period returns on seasoned bonds around the time of a city's new issue and general obligation debt burden of the city is significantly stronger for low frequency issuers than for high frequency issuers.

Since an interaction effect is expected, no hypotheses are posited for main effects. As discussed in the following section, the hypotheses are tested in a multiple regression model which controls for other potentially confounding factors.

3.3. Sample Selection, Model Development, and Testing Procedures

3.3.1. Sample Selection

To increase comparability with the Ingram et al. (1989) and the Ingram and Wilson (1999) studies, we use data encompassing the same time periods as those studies. We examine 594 new general obligation issues made by 197 cities from June 1978 to August 1989. These issues were selected from an initial sample of 1,105 new issues of 391 cities included in the Bond Buyer MUNIFICHE database. They represent essentially all advertised new general obligation issues made by cities over 25,000 population during the period, other than for refunding purposes. Of the 1,105 new issues, 84 were eliminated due to multiple issues on the same date, 395 were eliminated because no seasoned bond price data were available for those cities in our time-series bond price database,[8] and 32 were eliminated due to missing accounting or other data.

Although the final sample of 594 new issues was not randomly selected, it includes cities from 31 different states and every geographic region is well represented. Though the sample appears to be broadly representative of all cities over 25,000 population, the cities in the sample are larger on average than the total population and likely have better information disclosure. Thus, one should use due care in extrapolating our findings to the total population. Contemporaneous correlation due to event clustering is not considered a significant problem since the new issues are well distributed over time.

3.3.2. Model Development

Seasoned bond returns are expected to exhibit increased variability around the time of a new bond issue by the same municipality (Ingram et al., 1989). Market reaction is not expected more than one month prior to the month of sale, however, because of the prevailing low level of information dissemination by municipalities and the relatively thin market for municipal securities. For this reason, and to increase comparability with Ingram et al. (1989) we use a three-month "event" window – the month in which there is a new bond issue by the same municipality, the month preceding the event month, and the month following.

We use multiple regression analysis to examine the association between holding period returns (HPRs) on seasoned bonds and accounting measures, after controlling for other relevant factors. Because multiple factors affect investors' interpretations of information provided in new issue disclosures, it is not feasible to partition our sample into "good news" or "bad news" groups on the basis of accounting information alone. Rather, it is important to control for factors other than accounting information that might explain observed variation in returns surrounding a new issue.

One factor we control is the possibility that some seasoned bonds may be mispriced prior to a new issue by the same city because of insufficient entity-specific information. The rationale for controlling for this factor is that seasoned bonds with significant mispricing in the secondary market may simply adjust to the price of the new issue, rather than to any specific information included in the official statement for the new issue. If prices were available for the new issue, we would prefer to use a price difference variable (the price of the new issue minus the pre-new issue price of the seasoned bond) to control for possible mispricing. We would expect seasoned bonds for which there is a dearth of entity-specific information to be systematically under-priced (i.e. discounted for information risk). Such bonds would exhibit a positive price adjustment, and thus exhibit a positive return, at the time of a new issue by the same city.

Because we have available only net interest cost (NIC) for the new bond issue, a yield proxy, rather than prices, we include in our model the admittedly crude variable YDIF, defined as the difference between NIC on the new bonds and the average pre-new issue yield to maturity on the seasoned bonds over the three months $t-4$ through $t-2$, divided by the average pre-new issue yield. Because yields and prices move in opposite directions, YDIF is expected to be negatively associated with returns at the time of new issues.

Downgrades or upgrades of a city's general obligation bond rating in connection with a new issue provide a potentially strong signal of changed default risk. To control for the effect of rating changes we include two variables, DOWNCHG and UPCHG, indicating whether there was a rating downgrade or upgrade during

the month of, or the month preceding or following, a new issue.[9] We predict a negative association between DOWNCHG and returns and a positive association for UPCHG.

Given the twelve year period covered by our study, we must also control for at least three structural changes known to have occurred over this period. First, there was a shift in the market from mainly commercial bank and property casualty investors in the late 1970s and early 1980s to individual investors in the late 1980s (The Bond Market Association, 2001). Second, the Tax Reform Act (TRA) of 1986 had a strong effect on the municipal bond market; causing many issuers to rush into the market in 1985 before reduced tax rates reduced demand for tax-exempt securities. Third, the quality and reliability of government financial reports improved dramatically over the decade of the 1980s, particularly after creation of the GASB in 1984 and passage of the federal Single Audit Act of 1984. While it is not possible to predict the combined effect of these factors on bond returns, we include a single dummy variable AFT85 to indicate whether a new issue occurred before or after 1985. This partitioning of time is intended to capture the combined effects of the three changes.

Several types of risk affect the required yield and thus the price of a bond. Van Horne (1990) identifies these as interest rate risk, purchasing power (inflation) risk, call risk, and default risk. We control interest rate risk and purchasing power risk by adjusting returns by market returns for the same month. Call risk is not a problem because our sample includes only noncallable seasoned bonds. Although prior research (e.g. Duvall & Cheney, 1984; Van Horne, 1990; Weinstein, 1981) finds that default risk may influence price elasticity to some extent, default risk is primarily a function of duration of the bonds. Absent the data needed to compute duration, we include in our regression models the variable term-to-maturity, in natural logarithmic form (LOGMAT), to proxy for duration. We predict a positive association between returns and LOGMAT.

Finally, given that we examine new issues occurring over several years, the extent of market reaction to new issues may be a function of the level of market "risk averseness" existing at the time of the new issue. To control for this effect, we include the variable MRISK, measured as the spread between market yield rates for Moody's 20 year Aaa and Baa bond indexes for the same week as each new issue. No directional effect is predicted, however.

Based on our hypotheses, we expect generally stronger seasoned bond price reactions for infrequent issuers. The direction of association between frequency of issue and returns is indeterminate, however, since the composite signal conveyed by the new issue cannot be unambiguously interpreted. To control for the main effect of frequency of issue, we include the variable ISSFREQ indicating the number of new bond issues of each city during the sample period (1978–1989).

Theory provides little help in selecting accounting test variables. As Ingram et al. (1987) note, prior research examines many accounting-related variables, generally selected on an ad hoc basis. The accounting variables, the ratio of fund balance to revenue and the level of debt, are used by Ingram and Wilson (1999); therefore, for comparison purposes, they are also used in our analysis. Additionally, these ratios are known to be used by bond analysts to assess default risk and both are significant in prior studies (e.g. Raman & Wilson, 1990; Wallace, 1981; Wilson, 1983; Wilson & Howard, 1984). Although other accounting variables could have been included, collinearity is a concern. Finally, it should be noted that our focus is on the timing of informational effects on seasoned bond prices, not on fully explaining those effects.

Thus, we include FINPOS, defined as the ratio of fund balance of the general fund to revenues of the general fund. This variable is expected to be positively related to holding period returns, after controlling for other factors that affect price variability. That is, strong financial position should be interpreted favorably, and therefore, produce an increase in seasoned bond price. As hypothesized in H1, we expect financial position to convey more *new* information about the default risk of infrequent issuers than that of frequent issuers.

A second accounting variable, a measure of debt relative to the city's capacity to service the debt, is the ratio of general obligation debt to General Fund revenues (DEBT). This measure should be negatively associated with holding period returns. Again, the association is expected to be stronger for the bonds of infrequent issuers (H2).

Regarding the issue of levels vs. changes in the context of our two test variables, consider that the change in fund balance over revenues means little if an adequate fund balance still exists at the end of the period. In fact, a decrease in fund balance is considered good financial management if the beginning-of-period fund balance is considered excessively large. After all, the purpose of government is not to accumulate large surpluses. On the other hand, if secondary market investors have had no information about the financial position of the government for an extended period, they can be expected to react favorably, or at least not unfavorably, to the fact that the government currently has a sound financial position. In the absence of any news, the market could be expected to discount prices to hedge for the possibility of declining financial performance, thus increasing the probability of a positive reaction to strong financial position. Similar arguments can be made about changes in debt burden. Recall from earlier discussion that an increase (decrease) in debt is not automatically interpreted as bad (good) news. But, the market may react favorably to a low debt burden reported at the time of a new bond issue, or negatively to a high debt burden.

To test hypotheses H1 and H2 we utilize two test variables, FPOSFREQ and DEBTFREQ, formed by the interaction of the two variables FINPOS and DEBT

with a dichotomous variable FREQ, having a value of one if the frequency of new issues is less than the median number of new issues or zero if at or above the median. Significant coefficients for these variables (positive for FPOSFREQ and negative for DEBTFREQ) will support H1 and H2.

Based on the foregoing specification, the following regression model is estimated (subscripts for the *i*th bonds omitted for simplicity):

$$\begin{aligned}\text{HPR} = {} & b_0 + b_1\text{YDIF} + b_2\text{DOWNCHG} + b_3\text{UPCHG} + b_4\text{AFT85} \\ & + b_5\text{ISSFREQ} + b_6\text{LOGMAT} + b_7\text{MRISK} + b_8\text{FINPOS} \\ & + b_9\text{DEBT} + b_{10}\text{FPOSFREQ} + b_{11}\text{DEBTFREQ} + \mathbf{e}\end{aligned}$$

where HPR is the cumulative market-adjusted HPR for the three event month interval being examined, and all independent variables are as described in the preceding section and in Table 2. The ordinary least squares residual is **e**.

HPR is calculated as:

$$\text{HPR}_{it} = \frac{R_{it} - \text{MR}_t}{\text{SMR}_t}$$

where HPR_{it} is the risk adjusted holding period return for bond *i* and month *t*, R_{it} is the raw HPR for bond *i* and month *t*,[10] MR_t is the average market HPR for period *t*,[11] and SMR_t is the standard deviation of market HPRs for month *t*.

Subtracting MR_t from bond *i*'s return, R_{it}, removes market-wide sources of price change; whereas, deflating by a measure of market variance assigns less weight to return differences during periods of high market volatility. Failure to control for market volatility could induce spurious results, especially if market volatility differentially affects comparison and event period returns.

This model is estimated separately for the three-month interval centered on the new issue month and each of the two three-month non-event comparison intervals.[12]

Table 2 describes all variables used in the regression models.

3.4. Results

3.4.1. Descriptive Statistics

Table 3 presents descriptive statistics for the dependent and independent variables used in the regression models. As shown, holding period returns exhibit large cross-sectional variability. However, the variability is much more pronounced for the event (new issue) interval than for the non-event intervals.

Table 2. Description of Variables Used in the Regression of Holding Period Returns on Financial Position, Debt Burden and Control Variables.

Variable Name	Description
Dependent variables	
HPR-PRE	Holding Period Return, measured as the excess of return over the market return (measured as the average return for 324 seasoned bonds at time t), adjusted for market standard deviation $((R_{it} - \mathrm{MR}_t)/\mathrm{SMR}_t)$, cumulated over the pre-new issue event interval t_{-4} through period t_{-2}. Holding period return for month t is defined as $\log((P_0 + C)/P_{-1})$, where P is end-of-month price and C is the accrued monthly coupon payment.
HPR-NI	Same as above, except the adjusted holding period return is cumulated over the new issue event interval t_{-1} through period t_{+1}.
HPR-POST	Same as above, except the adjusted holding period return is cumulated over the post-new issue interval t_{+2} through period t_{+4}.
Independent variables	
Variables of interest	
FINPOS[a]	Current financial position measured as the ratio of General Fund fund balance to General Fund revenue, expressed as a percentage.
DEBT	The ratio of total general obligation debt outstanding to General Fund revenue.
FPOSFREQ	Interaction variable defined as FINPOS × FREQ, where FREQ is coded 1 for low frequency issuers (below the median frequency of 5 issues in 12 years) and 0 for high frequency issuers.
DEBTFREQ	Interaction variable defined as DEBT × FREQ, where FREQ is coded 1 for low frequency issuers (below the median) and 0 for high frequency issuers.
Control variables	
YDIF	Net interest cost on the new issue minus the average yield to maturity on the same entity's seasoned bond for the three-month pre-event interval, divided by the average pre-event yield.
DOWNCHG	Dummy variable indicating whether Moody's Investors Service downgraded the entity's general obligation bond rating during the three month new issue interval.
UPCHG	Dummy variable indicating whether Moody's Investors Service upgraded the entity's general obligation bond rating during the three month new issue interval.
AFT85	Dummy variable indicating whether the new issue occurred after 1985.
ISSFREQ	Number of general obligation bond issues of the municipality during the period 1978–1989 (excluding refunding issues).
LOGMAT	Natural logarithm of the term to maturity. (Source: Interactive Data Services, Inc.)
MRISK	Market default risk spread. Defined as the difference between Moody's Aaa and Baa yield indexes for same months as returns (Moody's Bond Survey, biweekly).

[a] The accounting variables were extracted from the most recent financial statement data provided in the bond prospectus for the new issue, generally the audited financial statements for the preceding fiscal year.

Table 3. Descriptive Statistics for a Sample of 594 General Obligation Bond Issues Made by 197 City Governments with Populations Greater than 25,000 Between 1978 and 1989 – Includes Variables Used in the Regression Model.

Variable Name[a]	Mean	Median	Standard Deviation	Minimum	Maximum
Dependent variables					
HPR-PRE	0.027	0.033	1.551	−7.388	8.409
HPR-NI	0.056	−0.030	2.294	−18.243	10.867
HPR-POST	−0.090	−0.071	1.357	−5.328	4.897
Independent variables					
YDIF	−0.038	−0.041	0.120	−0.460	0.447
DOWNCHG	0.024	0.000	0.152	0.000	1.000
UPCHG	0.035	0.000	0.185	0.000	1.000
AFT85	0.325	0.000	0.469	0.000	1.000
ISSFREQ	5.663	5.000	4.355	1.000	22.000
MAT	148.695	142.000	63.572	15.000	427.000
MRISK	0.965	0.983	0.410	0.313	1.900
FINPOS	13.616	10.913	12.856	−17.208	114.530
DEBT	1.425	1.122	1.290	0.041	13.900
FPOSFREQ	5.621	0.000	12.132	−10.875	114.530
DEBTFREQ	0.477	0.000	0.960	0.000	8.899

[a] Holding Period Returns (HPRs) are measured as the excess of return over the market return (measured as the average return for 324 seasoned bonds at time t), adjusted for market standard deviation $((R_{it} - \mathrm{MR}_t)/\mathrm{SMR}_t)$, cumulated over the event interval. HPRs are measured during the pre-new issue event interval t_{-4} through t_{-2} (HPR-PRE), the new issue event interval t_{-1} through t_{+1} (HPR-NI), and the post-new issue event interval t_{+2} through t_{+4} (HPR-POST). YDIF is the yield difference between the new issue and the average yield to maturity. DOWNCHG is a dummy variable indicating a bond rating downgrade during the three month new issue interval, while UPCHG is a dummy variable indicating a bond rating upgrade during the three month new issue interval. AFT85 indicates whether the issue occurred after 1985. ISSFREQ indicates the number of general obligation bond issues of the issuer in the period 1978–1989. Term-to-maturity (MAT), in months from the new issue month to the maturity date of the seasoned bond, is shown for descriptive purposes in months rather than the logarithmic form (LOGMAT) used in the regression models. MRISK is market default risk spread. FINPOS is the ratio of General Fund fund balance to General Fund revenue, while DEBT is the ratio of general obligation debt to General Fund revenue. FPOSFREQ and DEBTFREQ are the interactions of FINPOS and DEBT, respectively, with a dichotomous frequency of issue variable. For more complete definitions see Table 2.

3.4.2. Regression Results

Table 4 provides the regression results for the three-month new issue "event" interval as well as the pre- and post-new issue "non-event" intervals. As Table 4 shows, for the new issue interval model, three of the control variable coefficients are significant at the 0.01 level, with YDIF being strongly significant in the expected direction. A high negative value for YDIF indicates that the pre-new issue yield

Table 4. Regression Results for a Sample of 594 Seasoned Bonds, by City Governments Issuing New General Obligation Bonds Between 1978 and 1989, for Three Month New Issue Interval and Pre- and Post-New Issue Control Intervals – Dependent Variable is Cumulative Adjusted Holding Period Return.[a]

		Pre-New Issue Interval t_{-4}–t_{-2}		New Issue[d] Interval t_{-1}–t_{+1}		Post-New Issue Interval t_{+2}–t_{+4}	
Variable[b]	Pred.[c] Sign	Coef.	*t*-Stat.[e]	Coef.	*t*-Stat.	Coef.	*t*-Stat.
INTERCEPT	?	−3.320	4.12	−1.384	1.16	−0.864	1.12
YDIF	−	2.140	3.46***	−6.196	7.15***	−0.795	1.45
DOWNCHG	−	0.253	0.60	−1.499	2.48***	0.456	1.20
UPCHG	+	0.361	1.06	−0.089	0.18	−0.200	0.66
AFT85	?	0.266	1.49	0.756	2.87***	0.151	0.92
ISSFREQ	?	0.007	0.39	0.000	0.01	0.014	0.91
LOGMAT	?	0.722	4.64***	0.209	0.91	0.133	0.89
MRISK	?	−0.077	0.43	−0.021	0.08	−0.052	0.32
FINPOS	+	−0.012	1.52	−0.018	0.83	0.009	1.34*
DEBT	−	−0.046	0.74	0.096	1.08	−0.066	1.18
FPOSFREQ	+	0.011	1.15	0.027	2.07**	−0.009	1.08
DEBTFREQ	−	0.020	0.21	−0.354	2.63***	0.049	0.59
Adj. *R*-Sq.		0.042		0.113		−0.002	
Model *F* value		3.34***		7.83***		0.89	

[a] Holding Period Returns (HPRs) are measured as the excess of return over the market return (measured as the average return for 324 seasoned bonds at time *t*), adjusted for market standard deviation ($(R_{it} - \mathrm{MR}_t)/\mathrm{SMR}_t$), cumulated over the event interval.

[b] YDIF is the yield difference between the new issue and the average yield to maturity. DOWNCHG is a dummy variable indicating a bond rating downgrade during the three month new issue interval, while UPCHG is a dummy variable indicating a bond rating upgrade during the three month new issue interval. AFT85 indicates whether the issue occurred after 1985. ISSFREQ indicates the number of general obligation bond issues of the issuer in the period 1978–1989. Term-to-maturity, measured from the middle month of each interval to the maturity of the seasoned bond, is shown in logarithmic form (LOGMAT). MRISK is market default risk spread. FINPOS is the ratio of General Fund fund balance to General Fund revenue, while DEBT is the ratio of general obligation debt to General Fund revenue. FPOSFREQ and DEBTFREQ are the interactions of FINPOS and DEBT, respectively, with a dichotomous frequency of issue variable. For more complete definitions see Table 2.

[c] The predicted sign refers to coefficients for the new issue interval.

[d] Accounting measures are predicted to be significantly associated with the holding period return during the three month new issue interval but not during the pre-and post-new issue three month non-new issue intervals.

[e] Two-tailed unless + or − sign indicated.

*$p < 0.10$.

**$p < 0.05$.

***$p < 0.01$.

on the seasoned bond is higher than that of the new bond issue. Considering the inverse relationship between yields and prices, this suggests that initially underpriced or over-yielded seasoned bonds will increase in price and decrease in yield at the time of the new issue (i.e. exhibit a positive HPR). Thus, the relationship will be inverse between YDIF and HPR. The strong association for YDIF during the new issue interval is consistent with a presumption of initial mispricing and subsequent adjustment of seasoned bond prices to bring them in line with new issue pricing. As such, YDIF serves as a control for any omitted factors that may explain the initial difference between new issue prices and seasoned bond prices.[13]

Consistent with Marquette and Wilson (1992), returns during the new issue interval are also significantly associated with downgrades of bond ratings (DOWNCHG); whereas, upgrades appear to have no effect on returns. Although we do not predict direction, the significant finding for AFT85 indicates that structural factors affect market reaction to new issue information. While we were unable to identify which structural factors are contributing to the significance of AFT85, the significance does indicate the potential need for future studies to consider the effect of improved accounting/audit quality engendered by GASB and the Single Audit Act.

Interestingly, the results for the pre-new issue model suggest that seasoned bond returns appear to reflect only the logarithm of term-to-maturity (LOGMAT) and the yield difference (YDIF). The significant positive association for YDIF is likely because bonds with high pre-new issue positive returns would have had declining yields, which would have contributed to higher YDIFS. Results for the post-new issue model indicate an almost complete absence of new information being impounded in prices, indicating that prices have adjusted completely during the new issue event interval.

The variables of primary interest are the interaction of the two accounting variables FINPOS and DEBT with frequency of issue, FPOSFREQ and DEBTFREQ. Not surprisingly, there is no main effect observed for the full sample of seasoned bonds for FINPOS and DEBT. However, the significant associations for FPOSFREQ ($p < 0.05$) and DEBTFREQ ($p < 0.01$) support hypotheses H1 and H2, suggesting that accounting variables are more strongly associated with seasoned bond returns of infrequent issuers. Except for a weak main effect association for FINPOS in the post-new issue model, the results indicate that the secondary market only becomes aware of accounting information around the time of new issues.[14]

3.4.3. Sensitivity Analysis

In addition to the models with HPR as the dependent variable, we ran regressions in which the dependent variable was the natural logarithm of the price relative, LOG

(P_t/P_{t-1}). All results are virtually identical to those reported. We also tested an alternative dichotomous measure of FINPOS defined as adequate vs. inadequate fund balance, with FINPOS partitioned at 5%. The *t*-statistic for FPOSFREQ was 2.30 compared to 2.07 for the continuous measure.

Although we indicated earlier that contemporaneous correlation was not likely to be a problem since new issues were well distributed over time, there would still be a possibility of lack of independence over time for some issues, but mainly for the larger cities that issue bonds each year. To test for this effect, we ran the regression models using only one new issue for each city. Consistent with lower statistical power due to the smaller sample size, the statistical significance for FPOSFREQ was reduced from the 0.05 level to 0.10 and the significance for DEBTFREQ was reduced from the 0.01 level to 0.05.

We also conducted all tests with the event month defined alternatively as fiscal year end and month of annual report release. Consistent with Ingram et al. (1989), the accounting variables were not significantly associated with returns during the event interval for either of the two alternative events. Finally, we analyzed reduced samples of the data in which bonds were excluded if either the fiscal year end or annual report release month occurred within plus or minus one month of the new issue event month. The results for FPOSFREQ and DEBTFREQ were even stronger in each of these reduced samples than those reported in Table 4. This provides additional evidence that disclosures in the offering statements for new issues provide the information that primarily affects seasoned bond prices.

3.5. Conclusion

Using market-adjusted holding period returns, our results agree with the previous finding of Ingram et al. (1989) that seasoned bond prices exhibit strongly higher variability in the month of and, to a lesser extent, the month following a new issue. The issue of main concern, however, was whether accounting information is part of the total information set being used to adjust prices during the new issue interval. Our regression results indicate that accounting information is part of the total information impounded in seasoned bond prices during new issue intervals (contrary to Ingram & Wilson, 1999), but consistent with the market segmentation study of Feroz and Wilson (1992), accounting information appears to be value relevant in the secondary market only for the bonds of infrequent issuers.

4. SUMMARY

The purposes of this paper are threefold: (1) to provide an updated review of municipal capital markets research; (2) to extend empirical research involving

the secondary bond market; and (3) to provide direction for future research. The research reviewed and provided helps increase understanding of the relationship between accounting information and the municipal bond capital markets. The more recent research tends to support earlier research that finds an association between bond ratings and accounting information. Generally, the new issues market research indicates that the quality of accounting information and the relationship of accounting information to structural variables is important in new issues research. The research on how accounting information affects secondary market bond pricing is evolving, indicating that while accounting information is important to the secondary market, consideration needs to be given to the cost-benefit of obtaining accounting information and the information asymmetry that tends to exist in the secondary market.

While future research topics are discussed in each section there are some global topics and limitations to future research that remain to be addressed. No theory exists regarding how accounting information impacts the municipal bond market. As a result, the selection of financial accounting variables is made on an ad hoc basis. Because no theory identifies which financial accounting variables are most valuable to the capital markets it is difficult to tell whether studies adequately explain the significance of accounting information to the municipal bond markets.

Many of the studies cited rely on at least a portion of data from the pre-GASB, pre-Single Audit Act, pre-SEC disclosure implementation period. As a result, there continues to be a need for research that uses post-implementation data. The evidence in Section 3 (see Table 4) indicates that accounting-related factors, such as improved reliability of information, could be highly significant in the municipal bond market. Additionally, new GASB standards, such as Statement No. 34, add substantially to the amount and type of disclosure required, raising questions concerning the importance of the newly required disclosures to the government capital markets.

As shown, financial accounting information is useful in predicting bond ratings and bond rating changes. Researchers (Apostolou et al., 1985; Benson et al., 1991; Hefzi et al., 1988; Ingram & Wilson, 1999; Liu & Seyyed, 1991; Wallace, 1981; Wilson & Howard, 1984) also find that both financial accounting information and bond ratings are associated with net interest costs for new issues and yields for seasoned issues. Additionally, net interest costs are associated with the yield of seasoned issues (Ingram & Wilson, 1999). The relationships among the variables in governmental capital markets research raises statistical concerns. For studies that are looking at the association between financial accounting information and net interest cost or yields it would appear that models that omit a bond rating variable suffer from a potential omitted variables problem. However, since bond rating is an endogenous variable, using it as a right-hand variable may require the use of a simultaneous equation model (Kmenta, 1986).

NOTES

1. Two reasons for the dramatic improvement in governmental accounting and reporting are the creation of the Governmental Accounting Standards Board (GASB) in 1984 and the passage of the federal Single Audit Act in the same year. The GASB has issued standards that collectively have improved the consistency, comparability, relevance and reliability of state and local government financial reports. Further improving the quality and reliability of financial reports are the entity-wide audits required by Single Audit Act. More recently, the Securities and Exchange Commission (SEC) has worked to improve reporting disclosure in the municipal bond market.

2. One measure of the improvement in governmental financial reporting is the rapid growth in the number of municipalities receiving the Government Finance Officers Association's Certificate of Achievement for Excellence in Financial Reporting. Raman and Wilson (1994, p. 528) report, for example, that 55.7% of their sample of cities for the years 1984–1987 held the certificate, compared with only 18.5% of Feroz and Wilson's (1992, p. 488) sample of cities for the years of 1978–1983. Both samples are taken from cities with populations of 25,000 or greater.

3. The Municipal Securities Rule-Making Board (MSRB) has been actively engaged in establishing municipal bond issue standards since its inception in 1975. MSRB is a self-regulatory organization, designated by Congress to establish rules for the municipal securities market. In 1989, the SEC adopted rule 15c2–12, which basically requires that an official offering statement be made available at the time of a municipal offering. 15c2–12 was amended in 1994 to require that bond underwriters, brokers, and dealers ensure that the issuers whose municipal bonds they handle provide notification of material events and annual filings of financial information to one or more Nationally Recognized Municipal Securities Information Repositories (NRMSIRs). Note, however, that the SEC has no statutory authority to directly impose disclosure requirements on state and local governments that issue debt.

4. The Government Finance Officers Association (previously the Municipal Finance Officers Association) has sought to improve financial reporting since its inception in 1906 (Gauthier, 2001). Since 1945, it has awarded the Certificate of Conformance (Certificate of Achievement) for quality financial reporting (Gauthier, 2001). However, increased participation in the award program coincides with the formation of GASB and implementation of the Single Audit Act. Among other factors, the award criteria is based on compliance with generally accepted accounting principles, as promulgated by GASB.

5. The five items most often selected are general revenue to general expenditures, general obligation debt to population, percentage change in the population, short-term debt to revenue, and general revenue to general expenditures for prior year (Lewis et al., 1988).

6. The bankruptcy of Orange County in 1994 and the possible financial misrepresentation related to a Washington, DC bond issue in 1994 are two examples of such financial problems.

7. Default risk refers to the probability that the issuer will be unable to make principal and interest payments when due. Even though bond ratings provide a crude measure of default risk, Lamb and Rappaport (1987) note that ratings are only infrequently revised between new issues and normally are changed only when fairly major changes in default risk have occurred. Thus, investors can experience considerable economic gain or loss due

to price changes within the broad rating category assigned. Investors' perception of default risk may be affected by factors such as shifts in the local economy, and the city's financial position and political climate. Since many investors do not hold bonds to maturity, they stand to lose if prices decline because of increased default risk, even though actual defaults are infrequent.

8. These prices were available from Interactive Data Services (IDS) for a large number of bonds issued after 1977. As indicated earlier, proprietary pricing services, such as IDS are preferable since the prices are derived from different yield-to-maturity curves, corresponding to different levels of default risk, and a rating scale that assigns numerical values to a specific bond based upon the prices at which similar bonds are selling in the market.

9. In most bond price or yield studies, dummy variables for bond rating categories are included as a broad proxy for default risk. Because our model focuses on price changes (returns) rather than the level of prices or yields, rating changes are more relevant than rating categories.

10. For the sake of comparability, raw holding period returns for each bond are computed in the same manner as Ingram et al. (1989):

$$R = \ln(P_0 + C)/P_{-1}$$

where P_0 is the price at the end of month t_0, P_{-1} is the price at the end of month t_{-1}, and C is the accrued coupon payment for one month.

11. The market return is measured as the average return for 324 seasoned bond issues included in our time-series bond data base. Ingram (1985) utilizes several alternative index measures and finds that the results are not sensitive to the index used. Standard deviation of market returns is estimated by using the standard deviation of returns for all bonds for the same month as a new issue occurred. If an alternative market index, such as Moody's or Salomon Bros. is used, the variability of the bonds included in the index could not be determined.

12. A necessary precondition to testing the effects of the accounting test variables is to confirm that the variability of HPRs is greater in periods immediately surrounding new bond issues than in months where no such events have occurred. In substance, this test replicates the results of Ingram et al. (1989) to ensure their results hold with the data used in our study. We find, as expected, that the variability is sharply higher during the month of a new issue than for any of the months in the three-month announcement period, and that variability is significantly higher than expected during the three-month event period, but not in either of the two non-event comparison periods. Thus, our results support Ingram et al. (1989).

13. Although we believe YDIF is an important control variable, we also ran the regression with this variable omitted. The results for the test variables are considerably stronger. The *t*-ratios for FPOSFREQ increase from 2.07 to 2.59 and from 2.63 to 3.25 for DEBTFREQ.

14. Collinearity does not appear to be a problem in these regressions; the highest variance inflation factor is only 3.78 and that is for one of the interaction variables DEBTFREQ. White (1980) tests reveal no significant heteroscedasticity. To determine whether the results are affected by outliers, we omitted seven observations which are located in the upper tail of the residual and more than three studentized ranges away from the mean. The regression results are similar and slightly stronger with these observations deleted.

ACKNOWLEDGMENTS

The authors wish to thank Jere Francis, Robert Ingram, Inder Khurana, Jagan Krishnan, and workshop participants at the University of Alabama, DePaul University, and Temple University, and an anonymous reviewer for their helpful comments on prior versions of this paper. The financial support provided to Jacqueline Reck by the University of South Florida's Division of Sponsored Research is also gratefully acknowledged.

REFERENCES

Apostolou, N., Giroux, G., & Welker, R. (1985). The information content of municipal spending rate data. *Journal of Accounting Research*, *23*(2), 853–858.

Atiase, R. (1985). Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research*, *23*(Spring), 21–36.

Benson, E., Marks, B., & Raman, K. (1984). State regulation of accounting practices and municipal borrowing costs. *Journal of Accounting and Public Policy*, *3*(Summer), 107–122.

Benson, E., Marks, B., & Raman, K. (1986). The MFOA certificate of conformance and municipal borrowing costs. *Advances in Accounting*, *3*, 221–232.

Benson, E., Marks, B., & Raman, K. (1991). The effect of voluntary GAAP compliance and financial disclosure on governmental borrowing costs. *Journal of Accounting, Auditing & Finance*, *6*(3), 303–319.

The Bond Buyer Market Association (2001). *The fundamentals of municipal bonds*. New York: Wiley.

Bond Buyer Staff (2000). Are bonds underrated? The debate goes on. *GFOA Conference Daily* (13 June), *6*.

Chaney, B., & Copley, P. (2001). The association between accounting measures and interest cost on municipal bond issues: A test of entity relationships. Working Paper, North Carolina State University, Raleigh, North Carolina.

Copeland, R., & Ingram, R. (1982). The association between accounting information and municipal bond rating changes. *Journal of Accounting Research*, *20*(2), 275–289.

Copeland, R., & Ingram, R. (1983). Municipal bond market recognition of pension reporting practices. *Journal of Accounting and Public Policy*, *2*(Fall), 147–165.

Copeland, R., & Wilson, E. (1986). Managerial implications of municipal pension funding practices and bond borrowing costs. *Financial Accountability & Management*, *2*(2), 121–133.

Duvall, R., & Cheney, J. (1984). Bond beta and default risk. *Journal of Financial Research*, *6*(Fall), 243–254.

Fairchild, L., & Koch, T. (1998). The impact of state disclosure requirements on municipal yields. *National Tax Journal*, *51*(4), 733–753.

Feroz, E., & Wilson, E. (1992). Market segmentation and the association between municipal financial disclosure and net interest costs. *The Accounting Review*, *67*(July), 480–495.

Gauthier, S. (2001). *Governmental accounting, auditing, and financial reporting*. Chicago, IL: Government Finance Officers Association.

Governmental Accounting Standards Board (GASB) (1986). *Statement 5 of the Governmental Accounting Standards Board: Disclosure of pension information by public employee retirement systems and state and local governmental employers*. Norwalk, CT: Governmental Accounting Standards Board.

Governmental Accounting Standards Board (GASB) (1991). *Statement 14 of the Governmental Accounting Standards Board: The financial reporting entity*. Norwalk, CT: Governmental Accounting Standards Board.

Governmental Accounting Standards Board (GASB) (1994a). *Statement 25 of the Governmental Accounting Standards Board: Financial reporting for defined benefit pension plans and note disclosures for defined contribution plans*. Norwalk, CT: Governmental Accounting Standards Board.

Governmental Accounting Standards Board (GASB) (1994b). *Statement 26 of the Governmental Accounting Standards Board: Financial reporting for postemployment healthcare plans administered by defined benefit pension plans*. Norwalk, CT: Governmental Accounting Standards Board.

Governmental Accounting Standards Board (GASB) (1994c). *Statement 27 of the Governmental Accounting Standards Board: Accounting for pensions by state and local governmental employers*. Norwalk, CT: Governmental Accounting Standards Board.

Governmental Accounting Standards Board (GASB) (1999). *Statement 34 of the Governmental Accounting Standards Board: Basic financial statements and management's discussion and analysis-for state and local governments*. Norwalk, CT: Governmental Accounting Standards Board.

Hefzi, H., Ifflander, A., & Smith, D. (1988). Municipal bond market risk measures and bond ratings. *Advances in Accounting*, *6*, 111–126.

Icerman, R., & Welch, J. (1989). Data source and measurement differences in municipal accounting measures. *Journal of Accounting and Public Policy*, *8*, 143–160.

Ingram, R. (1983). The importance of state accounting practices for creditor decisions. *Journal of Accounting and Public Policy*, *2*(Spring), 5–17.

Ingram, R. (1985). A descriptive analysis of municipal bond price data for use in accounting research. *Journal of Accounting Research*, *23*(Autumn), 595–618.

Ingram, R. (1986). Tests of the fund accounting model for local governments. *Contemporary Accounting Research*, *3*(1), 200–221.

Ingram, R., Brooks, L., & Copeland, R. (1983). The information content of municipal bond rating changes: A note. *Journal of Finance*, *38*(June), 997–1003.

Ingram, R., & Copeland, R. (1982a). State mandated accounting, auditing and finance practices and municipal bond ratings. *Public Budgeting and Finance*, *2*(Spring), 19–29.

Ingram, R., & Copeland, R. (1982b). Municipal measures and reporting practices: An extension. *Journal of Accounting Research*, *20*(2), 766–772.

Ingram, R., & Copeland, R. (1986). *The effect of local government accounting practices on accounting numbers and creditor decisions*. Sarasota, FL: American Accounting Association.

Ingram, R., Raman, K., & Wilson, E. (1987). Governmental capital markets accounting research in accounting: A review. *Research in Governmental and Nonprofit Accounting*, *3*, 111–132.

Ingram, R., Raman, K., & Wilson, E. (1989). The information in governmental annual reports: A contemporaneous price reaction approach. *The Accounting Review*, *64*(2), 250–268.

Ingram, R., & Wilson, E. (1999). Information intermediation and seasoned municipal bond yields. *Research in Governmental and Nonprofit Accounting*, *10*, 1–31.

Kmenta, J. (1986). *Elements of econometrics*. New York: Macmillan.

Kotlikoff, L., & Smith, D. (1983). *Pensions in the American economy*. Chicago: University of Chicago Press and the National Bureau of Economic Research.

Lamb, R., & Rappaport, S. (1987). *Municipal bonds: The comprehensive review of tax exempt securities and public finance* (2nd ed.). New York: McGraw-Hill.

Langsam, S., & Kreuz, J. (1990). The relationship of type of audit opinion to governmental bond ratings. *Arkansas Business & Economic Review*, *23*(2), 1–10.

Lewis, B., Patton, J., & Green, S. (1988). The effects of information choice and information use on analysts' predictions of municipal bond rating changes. *The Accounting Review*, *63*(2), 270–282.

Liu, P., & Seyyed, F. (1991). The impact of socioeconomic variables and credit ratings on municipal bond risk premia. *Journal of Business Finance & Accounting*, *18*(5), 735–746.

Marks, B., & Raman, K. (1985). The importance of pension data for municipal and state creditor decisions: Replication and extensions. *Journal of Accounting Research*, *23*(2), 878–886.

Marquette, R., & Marquette, J. (1986). Improving models of municipal bond rating changes: Surrogation of subjective bias. *Research in Governmental and Nonprofit Accounting*, *2*, 235–259.

Marquette, R., & Wilson, E. R. (1992). The case for mandatory municipal disclosure: Do seasoned bond yields impound publicly available information? *Journal of Accounting and Public Policy*, *11*(Fall), 181–206.

McNichols, M., & Manegold, J. (1983). The effect of the information environment on the relationship between financial disclosure and security price variability. *Journal of Accounting and Economics*, *5*(April), 49–74.

Ohlson, J. (1979). On financial disclosure and the behavior of security prices. *Journal of Accounting and Economics*, *1*(3), 211–232.

Public Securities Association (1987). *Fundamentals of municipal bonds*. New York: Public Securities Association.

Raman, K. (1981). Financial reporting and municipal bond rating changes. *The Accounting Review*, *56*(4), 910–926.

Raman, K. (1982a). Financial reporting and municipal bond ratings. *Journal of Accounting, Auditing and Finance*, *5*(Winter), 144–153.

Raman, K. (1982b). Alternative accounting measures as predictors of municipal financial distress. *Journal of Accounting, Auditing and Finance*, *6*(Fall), 44–50.

Raman, K. (1986). Assessing the credit risk on municipal short-term debt. *Advances in Accounting*, *3*, 171–180.

Raman, K., & Wilson, E. R. (1990). The debt equivalence of unfunded governmental pension obligations. *Journal of Accounting and Public Policy*, *9*(Spring), 37–56.

Raman, K., & Wilson, E. R. (1994). Governmental audit procurement practices and seasoned bond prices. *The Accounting Review*, *69*(4), 517–538.

Reeve, J., & Herring, H. (1986). An examination of nonrated municipal bonds. *Journal of Economics and Business*, *38*(1), 65–76.

Securities and Exchange Commission (SEC) (1994). Municipal securities disclosure: Final rule. *Federal Register*, *59*(221), 59590–59610.

Sneed, C., & Sneed, J. (1999). The differential effect of state government pension funding practices on bond yields across varying maturities. *Research in Governmental and Nonprofit Accounting*, *10*, 33–47.

Soybel, V. (1992). Municipal financial reporting and the general obligation bond market: New York City, 1961–1975. *Journal of Accounting and Public Policy*, *11*, 207–231.

Stover, R. (1991). Bond ratings and new issue municipal bond pricing: Path analysis results. *Quarterly Journal of Business & Economics*, *30*(1), 3–17.

Van Horne, J. (1990). *Financial market rates and flows*. Englewood Cliffs, NJ: Prentice-Hall.

Vijayakumar, J. (1995). An empirical analysis of the factors influencing call decisions of local government bonds. *Journal of Accounting and Public Policy*, *14*, 203–231.

Wallace, W. (1981). The association between municipal market measures and selected financial reporting practices. *Journal of Accounting Research*, *19*(Autumn), 505–521.

Weinstein, M. (1981). The systematic risk of corporate bonds. *Journal of Quantitative and Financial Analysis*, *16*(September), 257–278.

Wescott, S. (1984). Accounting numbers and socioeconomic variables as predictors of municipal general obligation bond ratings. *Journal of Accounting Research*, *22*(Spring), 412–423.

Wescott, S. (1988). An assessment of the ability of accounting numbers to predict municipal general obligation bond ratings. *Research in Governmental and Nonprofit Accounting*, *4*, 79–102.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*, 817–838.

Wilson, E. (1983). Fiscal performance and municipal bond borrowing costs. *Public Budgeting and Finance*, *3*(Winter), 28–41.

Wilson, E., & Howard, T. (1984). The association between municipal market measures and selected financial reporting practices: Additional evidence. *Journal of Accounting Research*, *22*(Spring), 207–224.

Wilson, E., & Stewart, J. (1990). Municipal financial reporting and competition among underwriters for new issues of general obligation bonds. *Contemporary Accounting Research*, *6*(2), 573–592.

BOND INSURANCE AND GOVERNMENTAL ACCOUNTING RESEARCH

Earl D. Benson and Barry R. Marks

ABSTRACT

Prior research has assumed that accounting and auditing information has an impact on the borrowing costs of insured municipal bonds and that the default risk premiums on insured bonds differed by a fixed amount from the default risk premiums on uninsured bonds. We investigate these two hypotheses by analyzing the default risk premiums on insured and uninsured general obligation bonds for two different samples, school district bonds and city bonds. The findings suggest that the interest cost on both types of insured bonds is related to the underlying credit ratings associated with the issuer. However, the default risk premiums on the insured bonds are small in magnitude. Furthermore, the default risk premiums on the insured bonds do not behave in the same manner as the default risk premiums on uninsured bonds. The empirical results reject the hypothesis that the difference in default risk premiums between insured and uninsured general obligation bonds is a fixed amount. Because of the behavior and size of the default risk premiums on the insured bonds, we recommend that researchers focus their attention on analyzing the interest cost on uninsured bonds when doing studies on the impact of accounting and auditing information on issuer interest cost.

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 35–50

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11002-1

1. INTRODUCTION

Our objective is to compare the default risk premiums on insured bonds with the default risk premiums on uninsured bonds and to discuss the implications of these findings for governmental accounting research. Prior research on the relationships between municipal borrowing costs and governmental accounting and auditing information has not examined this issue closely. Just a few municipal bonds had insurance at the time that most early studies were conducted. However, Smith (2001) reports that nearly half of all municipal bonds issued in the first-half of 2001 had bond insurance.

Wallace (1981) assumed that the interest cost on insured bonds was dependent upon the characteristics of the bond issuer and the bond insurer. Wallace included a zero-one dummy variable to indicate the presence of bond insurance in her regression model of the relationship between accounting and auditing variables and the net interest cost on new bond issues. Some of her models included dummy variables for each of Moody's bond ratings and a dummy variable for bond insurance. Wilson and Howard (1984) replicated Wallace's study on a larger sample of bonds. They modeled the impact of insurance on municipal borrowing cost in the same manner as Wallace. At that time Moody's published the credit rating for each bond ignoring the impact of any insurance. Moody's changed its policy in 1984. An insured bond is now rated Aaa. Moody's still evaluates the creditworthiness of each insured bond issuer. If the municipality has some uninsured bonds outstanding, then Moody's will publish the credit rating for these bonds which effectively represents the underlying credit rating for the entity. Otherwise, Moody's will not publish the underlying credit rating of the entity unless requested by the issuer.

For this study default premiums are estimated for both city general obligation bonds and school district general obligation bonds. The measure of default risk is Moody's bond rating. While previous research has shown that default risk premiums exist for uninsured bonds, we are unaware of any analysis of default risk premiums for insured bonds since Moody's instituted its policy in 1984 of classifying all insured bonds as Aaa. If the default risk premiums on insured bonds are not statistically significant, then the capital markets do not consider the credit worthiness of the issuer in the pricing these bonds. Hence, any study investigating the relationship between municipal borrowing costs and the issuer's accounting, auditing, or credit rating information should exclude any insured bonds from the sample. However, if the default risk premiums on insured bonds are statistically significant, the capital markets examine the credit worthiness of the insurance provider and the credit worthiness of the issuer. Wallace (1981) and Wilson and Howard (1984) implicitly assume that the difference in default risk premiums on insured bonds and uninsured bonds is constant across all bond

rating categories. If the difference in default risk premiums is not constant across all bond rating categories, a different approach is necessary for modeling default risk in samples that contain insured and uninsured bonds.

Section 2 describes the operations of a private bond insurance company, Ambac Assurance Corporation, one of the four largest private bond insurance companies (Smith, 2001). Section 3 describes the model and data. Section 4 contains the empirical results, while Section 5 discusses the implications of the results for future governmental accounting research.

2. PRIVATE BOND INSURANCE

The focus in this section is on one private bond insurer, Ambac. While the insured bonds used in this study were insured by four different insurers, including FGIC, MBIA, and FSA, the information presented here about Ambac is representative of all four. Ambac Financial Group, Inc. is a holding company whose subsidiaries provide guarantee insurance and financial management services (Ambac Financial Group, Inc. 10-K Report, 1997). Ambac Financial Group, Inc. issues guarantee insurance through its subsidiary, Ambac Assurance Corporation. Ambac Assurance Corporation started the private municipal bond insurance industry in 1971. Ambac Assurance Corporation guarantees timely payment of interest and principal payments not only on municipal bonds but also structured finance and asset-backed obligations. Structured finance and asset-backed obligations are debt issues secured by a specific pool of assets, such as home mortgages, home equity loans, credit card receivables, or trade receivables. Ambac Assurance Corporation insures such obligations both in the U.S. and international markets.

Ambac Assurance Corporation is subject to the insurance laws and regulations of the states and countries where it conducts its business. The insurance regulatory commissions of these states and countries have the right to undertake audits to ensure that Ambac Assurance Corporation is in compliance with applicable laws and regulations. Its 1997 10-K Report specifically mentions the laws of the State of New York. The New York Financial Guaranty Insurance Corporations Law requires that financial guaranty insurance corporations have at least $2,500,000 in paid-in capital and to maintain a minimum surplus to policyholders of at least $65,000,000. The corporations must establish a contingency reserve. For municipal general obligation bonds, the corporation must make contributions each year to a contingency reserve equal to one-twentieth of the greater of 50% of premiums written or 0.55% of principal guaranteed. The corporations must maintain a surplus to policyholders and contingency reserves of no less than one-three-hundredth of its aggregate net liability for insured municipal bonds. The aggregate net liability is defined to be the guaranteed principal and interest payments less reinsurance and

collateral. The limit on the maximum amount that can be guaranteed for any one entity is tied to the entity's average annual debt service. An entity's average annual debt service, net of reinsurance and collateral, shall not exceed 10% of the corporation's surplus to policyholders plus contingency reserve. Furthermore, the entity's insured unpaid principal, net of reinsurance and collateral, cannot be more than 75% of the corporation's surplus to policyholders plus contingency reserve.

The insurance regulations placed upon private insurance companies should provide sufficient security to enable the companies to make timely payment of interest and principal on defaulted bonds in most circumstances. Hence, in 1984, Moody's decided to rate the insured bonds of the four major private insurance companies as "Aaa." However, during a deep recession or depression, these private insurance companies may not have enough resources to make timely payment of interest and principal on insured bonds.

This study examines both school districts and cities. During the depression, school districts had a much lower default rate than incorporated municipalities. In 1935, the percentage of school districts in default was 0.49%, while the percentage of incorporated municipalities in default was 5.20% (Hillman, 1936). As Hillman states "The bond man's contention that school bonds are one of the safest municipals seems borne out by the above statistics."

Another difference between cities and school districts is the diversity of their revenue sources. The primary sources of revenue for school districts are state aid and property taxes. For fiscal years ending in 1997, public schools received 48.8% of their revenue from state sources and 44.8% from local sources (U.S. Census Bureau, 2000b). Property taxes represented 65.6% of the revenue received by public schools from local sources.

Cities and townships, on the other hand, receive only 20.7% of their general revenue from the state and 71.7% of their general revenue from local sources (U.S. Census Bureau, 2000a). They are much less dependent upon state aid than school districts. Also, they are not as dependent upon property taxes because they can collect additional funds through sales taxes, income taxes, and charges for services. Property taxes are 28.9% of revenue from local sources. Because of the different historical default rates for school districts compared with cities and the different revenue structures, the default risk premiums may not be the same for school district bonds and city bonds.

3. MODEL AND DATA

The empirical model analyzed in this paper contains variables included in previous studies (Benson, 1979, 1999; Cook, 1982). The variables reflect market conditions

at the time of sale, the individual characteristics of the bonds, and default risk. The linear regression model can be stated as the following:

$$\text{TIC} = f(\overset{+}{\text{Dbb40}}, \overset{-}{\text{Lnsize}}, \overset{+}{\text{Dur}}, \overset{-}{\text{Lnbids}}, \overset{-}{\text{Ref}}, \overset{+}{\text{Call}}, \overset{-}{\text{Cprem}}, \overset{-}{\text{Bq}}, \overset{+}{A}, \overset{+}{\text{Baal}}, \overset{+}{\text{Baa}})$$

where the sign above the variable name in the equation is the anticipated sign of that variable's coefficient. The variables in the model are the following:

TIC = the true interest cost of the bond issue;
Dbb40 = the general level of municipal interest rates on the day before the sale as measured by the Bond Buyer 40 Index;
Lnsize = is the natural log of the dollar size of the issue;
Dur = is the duration of the bond issue;
Lnbids = the logarithm of the number of bids for the bond issue;
Ref = 1 if the bond is a refunding bond and otherwise 0;
Call = 1 if the bond is callable and otherwise 0;
Cprem = call premium for the callable bonds;
Bq = 1 if the issue is bank-qualified and otherwise 0;
A = 1 if the underlying credit rating for the entity is *A* by Moody's for insured bonds or the bond's credit rating is *A* by Moody's for uninsured bonds and otherwise 0;
Baa1 = 1 if the underlying credit rating for the entity is Baa1 by Moody's for insured bonds or the bond's credit rating is Baa1 by Moody's for uninsured bonds and otherwise 0;
Baa = 1 if the underlying credit rating for the entity is Baa by Moody's for insured bonds or the bond's credit rating is Baa by Moody's for uninsured bonds and otherwise 0.

TIC is a measure of the interest cost of the issue. If the measure of interest cost for the new bond issue reported in *The Bond Buyer* is NIC, net interest cost, then TIC was calculated for the winning bid using the procedure described in Benson (1999). TIC takes into consideration the time value of money, while NIC does not. TIC is computed from the proceeds that the bond issuer receives from the bond underwriters. TIC does not include the bond insurance premium or the cost of obtaining the bond rating. Bond insurance premiums are private information that is not reported in *The Bond Buyer*. The default risk measure for insured bonds is the underlying credit rating of the entity, while the default risk measure for uninsured bonds is the bond's credit rating. Since all the bonds in this study are general obligation bonds, these default measures are equivalent. Earlier models used the log of the average maturity as an independent variable in the model. Benson (1999) computed duration from the data from *The Bond Buyer*. Duration is a better measure

of interest rate risk because its computation is dependent upon the timing of the interest and principal payments. Fabozzi (2000) shows how duration is correlated with the change in a bond's price for a given change in interest rates. The other variables are defined as in most municipal bond studies.

All the variables except for the underlying credit ratings are either directly taken from *The Bond Buyer* or calculated from its data. The sample data were from the period June 1992 – July 1996. This included 593 school district and 964 city bonds. The underlying credit ratings were collected from *Moody's Bond Record*. Insured bonds were included in the sample only if the issuer's underlying credit rating was reported in *Moody's Bond Record*. The underlying credit ratings for the insured bonds were primarily A1, *A*, Baa1, or Baa. Since only a few insured bonds had an underlying credit rating above A1, these bonds were not included in the sample. Insured bonds included those insured by Ambac, FGIC, MBIA, and FSA. In order to analyze a comparable sample for uninsured bonds, the sample of uninsured bonds was limited to bonds rated A1, *A*, Baa1, or Baa. The variable A1 is the omitted dummy variable in the regression model.

Previous studies (Copeland & Ingram, 1982; Raman, 1981; Wallace, 1981; Wilson & Howard, 1984) demonstrate that bond ratings are related to the economic, accounting, and auditing characteristics of the issuer. Ingram and Wilson (1999) conclude that accounting information is impounded in bond ratings and bond prices primarily at the time that new bond issues are sold. The underlying credit rating used in this study for each insured bond issue was the rating that existed at the time of its sale.

Table 1 contains descriptive statistics for the 337 insured and the 256 uninsured school district bonds, while Table 2 contains descriptive statistics for the 381 insured and the 583 uninsured city bonds. The insured school district bonds have a longer duration and higher call premiums than uninsured school district bonds. Uninsured school district bonds are more likely to be bank-qualified than insured school district bonds. These same patterns are observed in comparing insured city bonds with uninsured city bonds.

Six different regression equations are reported. The previously described regression equation is estimated for each of our four samples, insured school district bonds, uninsured school district bonds, insured city bonds, and uninsured city bonds. A modified regression equation described below is estimated for two pooled samples: (1) insured and uninsured school district bonds; and (2) insured and uninsured city bonds. These results will be utilized to examine the following two hypotheses:

H1. The bond market is not using the characteristics of the individual bond issuer in pricing insured bonds.

Table 1. Descriptive Statistics for a Sample of 337 Insured School District General Obligation Bonds and a Sample of 256 Uninsured School District General Obligation Bonds Issued between June 1992 and July 1996.

Panel A. Continuous Variables

Continuous Variables[a]	Insured School District Bonds Number of Observations 337				Uninsured School District Bonds Number of Observations 256			
	Mean	Standard Deviation	Minimum	Maximum	Mean	Standard Deviation	Minimum	Maximum
TIC	5.347	0.462	4.113	6.697	5.028	0.655	3.361	6.896
Dbb40	6.127	0.372	5.350	7.370	6.108	0.389	5.390	7.370
Lnsize	15.678	0.766	13.816	19.832	15.066	0.739	13.816	17.316
Dur	7.353	1.668	2.807	11.653	6.218	1.941	2.157	11.613
Lnbids	1.323	0.428	0.000	2.485	1.487	0.542	0.000	3.091
Cprem	0.743	0.921	0.000	2.000	0.273	0.655	0.000	2.000

Panel B. Binary Variables

Binary Variables[a]	Insured School District Bonds Number of Observations 337	Uninsured School District Bonds Number of Observations 256
	Number of Nonzero Observations	Number of Nonzero Observations
Ref	14	18
Call	237	147
Bq	216	221
A1	65	65
A	176	143
Baa1	66	20
Baa	30	28

[a] Variable Definitions

(a) TIC is the true interest cost of the bond issue.
(b) Dbb40 is the general level of municipal interest rates on the day before the sale as measured by the Bond Buyer 40 Index.
(c) Lnsize is the natural log of the dollar size of the issue.
(d) Dur is the duration of the bond issue.
(e) Lnbids is the logarithm of the number of bids for the bond issue.
(f) Ref is 1 if the bond is a refunding bond and otherwise 0.
(g) Call is 1 if the bond is callable and otherwise 0.
(h) Cprem is the call premium for the callable bonds.
(i) Bq is 1 if the issue is bank-qualified and otherwise 0.
(j) A1 is 1 if the underlying credit rating for the school district is A1 by Moody's and otherwise 0. A is 1 if the underlying credit rating for the school district is *A* by Moody's and otherwise 0. Baa1 is 1 if the underlying credit rating for the school district is Baa1 by Moody's and otherwise 0. Baa is 1 if the underlying credit rating for the school district is Baa by Moody's and otherwise 0.

Table 2. Descriptive Statistics for a Sample of 381 Insured City General Obligation Bonds and a Sample of 583 Uninsured City General Obligation Bonds Issued between June 1992 and July 1996.

Panel A. Continuous Variables

Continuous Variables[a]	Insured City Bonds Number of Observations 381				Uninsured City Bonds Number of Observations 583			
	Mean	Standard Deviation	Minimum	Maximum	Mean	Standard Deviation	Minimum	Maximum
TIC	5.368	0.467	4.132	6.996	5.091	0.558	3.516	6.752
Dbb40	6.144	0.384	5.350	7.330	6.108	0.377	5.350	7.370
Lnsize	15.519	0.833	13.834	18.471	14.784	0.716	13.816	17.499
Dur	7.187	1.442	2.686	11.830	6.360	1.586	1.719	11.772
Lnbids	1.342	0.394	0.000	2.485	1.619	0.509	0.000	2.773
Cprem	1.125	0.970	0.000	3.500	0.509	0.858	0.000	3.000

Panel B. Binary Variables

Binary Variables[a]	Insured City Bonds Number of Observations 381	Uninsured City Bonds Number of Observations 583
	Number of Nonzero Observations	Number of Nonzero Observations
Ref	19	73
Call	318	454
Bq	258	507
A1	114	233
A	154	264
Baa1	79	50
Baa	34	36

[a] Variable Definitions

(a) TIC is the true interest cost of the bond issue.
(b) Dbb40 is the general level of municipal interest rates on the day before the sale as measured by the Bond Buyer 40 Index.
(c) Lnsize is the natural log of the dollar size of the issue.
(d) Dur is the duration of the bond issue.
(e) Lnbids is the logarithm of the number of bids for the bond issue.
(f) Ref is 1 if the bond is a refunding bond and otherwise 0.
(g) Call is 1 if the bond is callable and otherwise 0.
(h) Cprem is the call premium for the callable bonds.
(i) Bq is 1 if the issue is bank-qualified and otherwise 0.
(j) A1 is 1 if the underlying credit rating for the city is A1 by Moody's and otherwise 0. *A* is 1 if the underlying credit rating for the city is *A* by Moody's and otherwise 0. Baa1 is 1 if the underlying credit rating for the city is Baa1 by Moody's and otherwise 0. Baa is 1 if the underlying credit rating for the city is Baa by Moody's and otherwise 0.

H2. The difference in default risk premiums between insured and uninsured bonds is constant across bond ratings.

The first hypothesis is examined by testing if each of the coefficients for the underlying credit rating variables is not significantly different from zero in the two regression equations for insured bonds. If all the coefficients are not statistically different from zero, then the capital markets are not considering the default risk of the issuer of the insured bonds in pricing the bonds.

The second hypothesis is first investigated by comparing the default risk premiums for insured bonds with the default risk premiums for uninsured bonds. The default risk premiums are the coefficients of the credit rating variables in each of the regression equations. A regression model that contains insured and uninsured bonds is utilized to formally test the second hypothesis. The new regression model includes all the variables from the preceding regression model plus the following variables:

Ins = 1 if the bond is insured and otherwise 0;
Insa = 1 if the bond is insured and its underlying credit rating is *A* and otherwise 0;
Insbaa1 = 1 if the bond is insured and its underlying credit rating is Baa1 and otherwise 0; and
Insbaa = 1 if the bond is insured and its underlying credit rating is Baa and otherwise 0.

The variable Ins allows insured bonds to have a fixed difference in TIC across all bond ratings. By adding the interaction variables Insa, Insbaa1, and Insbaa to the regression equation, the difference in default risk premiums between an insured bond with an A underlying credit rating and an uninsured bond with an A credit rating becomes the coefficient of variable Ins plus the coefficient of variable Insa. If any of the coefficients of Insa, Insbaa1, or Insbaa are statistically significant, then the second hypothesis can be rejected because the default risk premiums for insured bonds are different than the default risk premiums for uninsured bonds by more than a fixed amount.

4. EMPIRICAL RESULTS

The ordinary least-squares regression equations for the insured school district bonds and uninsured school district bonds appear in Table 3. The variables that are statistically significant have the anticipated sign. The R^2 for insured school district regression is 0.925, while the R^2 for uninsured school district regression is 0.895.

Table 3. Regression Results with True Interest Cost as the Dependent Variable for a Sample of 337 Insured School District General Obligation Bonds and a Sample of 256 Uninsured School District General Obligation Bonds Issued between June 1992 and July 1996.

Variable[a]	Insured School District Bonds			Uninsured School District Bonds		
	Coefficient	*t*-Statistic	VIF	Coefficient	*t*-Statistic	VIF
Intercept	−2.218	−9.10***	0	−2.766	−6.57***	0
Dbb40	1.042	53.90***	1.049	1.219	33.55***	1.083
Lnsize	−0.017	−1.34	1.918	−0.036	−1.68*	1.399
Dur	0.132	24.73***	1.606	0.169	18.28***	1.745
Lnbids	−0.101	−5.53***	1.238	−0.092	−3.03***	1.461
Ref	0.007	0.18	1.047	−0.143	−2.60***	1.075
Call	0.193	8.41***	2.240	0.111	2.94***	1.911
Cprem	−0.041	−4.19***	1.620	0.003	0.13	1.251
Bq	−0.103	−5.48***	1.669	−0.161	−3.50***	1.365
A	0.043	2.28**	1.833	0.015	0.43	1.658
Baa1	0.041	1.77*	1.755	0.197	3.27***	1.428
Baa	0.108	3.59***	1.490	0.414	7.69***	1.540
Number of observations	337			256		
R^2	0.925			0.895		

Note: The *t*-statistic with one degree of freedom tests that the coefficient is equal to zero. VIF is the variance inflation factor, which is a measure of multicollinearity.

[a] The variables are defined in Table 1.

* Significant at the 0.10 level.

** Significant at the 0.05 level.

*** Significant at the 0.01 level.

The Variance Inflation Factor, VIF, is a measure of multicollinearity. The largest VIF in the two regression equations is 2.240. According to Gunst and Mason (1980), multicollinearity is not a problem when the VIFs are below 10.0. The *t*-statistic tests for the significance of the coefficient of an independent variable in an ordinary least squares regression equation. To check for the possible effect of heteroscedasticity on our statistical tests, we also computed White's test statistic to verify the statistical significance of each coefficient in the regression equations (White, 1980). The statistical results using White's procedure did not change the statistical significance of any of the variables in Table 3. The Shapiro-Wilk statistic does not reject the null hypothesis at the 0.05 level that the residuals are normally distributed in both regression equations.

The coefficients of *A*, Baa1, and Baa in the insured school district regression equation (Table 3) are 0.043, 0.041, and 0.108. The coefficient of 0.108 indicates

that the TIC is 0.108 percentage points higher for an insured school district bond with an underlying credit rating of Baa compared to an insured school district bond with an underlying credit rating of A1. Each of these variables is statistically significant at the 0.10 level. Hence, the first hypothesis that the bond market is not using the characteristics of the individual bond issuer in pricing insured bonds is rejected. The coefficients of *A*, Baa1, and Baa in the uninsured school district regression equation (Table 3) are 0.015, 0.197, and 0.414. The size of these coefficients is increasing at a much faster rate across bond ratings compared to the coefficients of the insured bonds. Such a pattern is inconsistent with the second hypothesis that the difference in default risk premiums is constant across bond rating categories for both insured and uninsured bonds.

Table 4. Regression Results with True Interest Cost as the Dependent Variable for the Pooled Sample of Insured and Uninsured School District General Obligation Bonds Issued between June 1992 and July 1996.

Variable[a]	Coefficient	*t*-Statistic	VIF
Intercept	−2.480	−10.36**	0
Dbb40	1.116	56.61**	1.046
Lnsize	−0.007	−0.59	1.898
Dur	0.151	29.09**	1.773
Lnbids	−0.107	−6.12**	1.351
Ref	−0.081	−2.44*	1.045
Call	0.164	7.46**	2.064
Cprem	−0.031	−2.96**	1.505
Bq	−0.133	−6.37**	1.595
A	−0.003	−0.11	3.513
Baa1	0.188	4.00**	5.150
Baa	0.400	9.47**	2.949
Ins	0.087	2.64**	4.991
Insa	0.051	1.35	5.577
Insbaa1	−0.143	−2.55*	5.879
Insbaa	−0.275	−4.74**	3.035
Number of observations	593		
R^2	0.907		

Note: The *t*-statistic with one degree of freedom tests that the coefficient is equal to zero. VIF is the variance inflation factor, which is a measure of multicollinearity.

[a] Variable Definitions: Ins is 1 if the bond is insured and otherwise 0. Insa is 1 if the bond is insured and its underlying credit rating is A and otherwise 0. Insbaa1 is 1 if the bond is insured and its underlying credit rating is Baa1 and otherwise 0. Insbaa is 1 if the bond is insured and its underlying credit rating is Baa and otherwise 0. The other variables are defined in Table 1.

* Significant at the 0.05 level.

** Significant at the 0.01 level.

The ordinary least-squares regression equation for the pooled sample of insured and uninsured school district bonds is located in Table 4. The coefficients of Insbaa1 and Insbaa are statistically significant with negative signs, which indicates that the default risk premiums for these insured bonds are not a fixed amount but vary according to rating and are different from uninsured bonds. So the second hypothesis is rejected. Although some of the VIFs are higher than in the previous regression equations, they are still below 10. The results do not change if White's test statistic is used instead of the *t*-statistic to determine the statistical significance of the coefficients. The Shapiro-Wilk statistic is not significant at the 0.05 level.

The ordinary least-squares regression equations for the insured city bonds and uninsured city bonds appear in Table 5. The variables that are statistically significant have the anticipated sign. The R^2 for the insured city bond regression is 0.928, while the R^2 for the uninsured city bond regression is 0.917. The largest

Table 5. Regression Results with True Interest Cost as the Dependent Variable for a Sample of 381 Insured City General Obligation Bonds and a Sample of 583 Uninsured City General Obligation Bonds Issued between June 1992 and July 1996.

Variable[a]	Insured City Bonds			Uninsured City Bonds		
	Coefficient	*t*-Statistic	VIF	Coefficient	*t*-Statistic	VIF
Intercept	−1.465	−7.09***	0	−2.276	−10.33***	0
Dbb40	1.067	61.26***	1.056	1.095	60.14***	1.040
Lnsize	−0.048	−4.51***	1.822	−0.015	−1.26	1.523
Dur	0.146	28.95***	1.256	0.181	35.09***	1.480
Lnbids	−0.022	−1.23	1.211	−0.143	−8.44***	1.634
Ref	−0.067	−2.16**	1.078	−0.107	−5.19***	1.036
Call	0.109	4.72***	1.746	0.073	3.56***	1.609
Cprem	−0.010	−1.23	1.595	−0.017	−1.79*	1.547
Bq	−0.133	−7.36***	1.699	−0.149	−6.40***	1.365
A	−0.015	−0.95	1.456	0.055	3.66***	1.237
Baa1	0.022	1.10	1.522	0.197	7.39***	1.235
Baa	0.072	2.79***	1.267	0.446	14.49***	1.212
Number of observations	381			583		
R^2	0.928			0.917		

Note: The *t*-statistic with one degree of freedom tests that the coefficient is equal to zero. VIF is the variance inflation factor, which is a measure of multicollinearity.

[a] The variables are defined in Table 2.

* Significant at the 0.10 level.

** Significant at the 0.05 level.

*** Significant at the 0.01 level.

VIF in the two regression equations is 1.822. We also computed White's test statistic to verify the statistical significance of each coefficient in the regression equations. The statistical results using White's procedure did not change the statistical significance of any of the variables in Table 5. The Shapiro-Wilk statistic does not reject the null hypothesis at the 0.05 level.

The coefficients of *A*, Baa1, and Baa in the insured city regression equation (Table 5) are −0.015, 0.022, and 0.072. Only the coefficient of Baa is statistically significant. Hence, the first hypothesis that the bond market is not using the characteristics of the individual bond issuer in pricing insured bonds is rejected. The coefficients of *A*, Baal, and Baa in the uninsured city regression equation (Table 5) are 0.055, 0.197, and 0.446. The size of these coefficients is very similar

Table 6. Regression Results with True Interest Cost as the Dependent Variable for the Pooled Sample of Insured and Uninsured City General Obligation Bonds Issued between June 1992 and July 1996.

Variable[a]	Coefficient	*t*-Statistic	VIF
Intercept	−2.070	−13.07**	0
Dbb40	1.085	81.38**	1.041
Lnsize	−0.025	−3.03**	1.995
Dur	0.170	44.96**	1.449
Lnbids	−0.110	−8.59**	1.575
Ref	−0.093	−5.35**	1.057
Call	0.079	4.98**	1.627
Cprem	−0.008	−1.23	1.670
Bq	−0.128	−8.33**	1.562
A	0.048	3.41**	1.977
Baa1	0.199	8.04**	2.894
Baa	0.461	16.12**	2.234
Ins	0.109	5.76**	3.456
Insa	−0.054	−2.28*	3.040
Insbaa1	−0.162	−4.87**	3.375
Insbaa	−0.379	−9.19**	2.353
Number of observations	964		
R^2	0.920		

Note: The *t*-statistic with one degree of freedom tests that the coefficient is equal to zero. VIF is the variance inflation factor, which is a measure of multicollinearity.

[a] Variable Definitions: Ins is 1 if the bond is insured and otherwise 0. Insa is 1 if the bond is insured and its underlying credit rating is *A* and otherwise 0. Insbaa1 is 1 if the bond is insured and its underlying credit rating is Baa1 and otherwise 0. Insbaa is 1 if the bond is insured and its underlying credit rating is Baa and otherwise 0. The other variables are defined in Table 2.

* Significant at the 0.05 level.

** Significant at the 0.01 level.

to those of the uninsured school districts. Again, the size of these coefficients is increasing at a much faster rate across bond ratings compared to the coefficients of the insured bonds. Such a pattern is inconsistent with the second hypothesis that the difference in default risk premiums is constant across bond rating categories for both insured and uninsured bonds.

The ordinary least-squares regression equation for the pooled sample of insured and uninsured city bonds is located in Table 6. The coefficients of Insa, Insbaa1, and Insbaa are statistically significant with negative signs, which indicates that the default risk premiums for these insured bonds are not a fixed amount but vary according to rating. So the second hypothesis is rejected. The highest VIF is 3.456. The results do not change if White's test statistic is used instead of the *t*-statistic to determine the statistical significance of the coefficients. The Shapiro-Wilk statistic is not significant at the 0.05 level.

5. CONCLUSION

These results have important implications for future research in governmental accounting and credit markets. The bond market does include the characteristics of the issuer in the pricing of insured school district and insured city bonds. Although the risk premiums on the insured bonds are statistically significant, their impact on the TIC of the bonds is quite small. Earlier models of the impact of accounting and auditing information on the municipal borrowing costs assumed that bond insurance changed borrowing costs by a fixed amount across all bond ratings. Our results indicate that such an assumption is not correct. Default risk premiums on uninsured bonds increase rapidly as the bond rating drops from A rated bonds to Baa1 and Baa rated bonds, while the default risk premiums on insured bonds increase at a slower rate as the rating drops. Since the default risk premiums on insured bonds are so small, researchers will have difficulty finding any relationship between municipal borrowing costs and accounting and auditing information for insured bonds unless they have a large number of Baa bonds in their sample. So researchers should focus on analyzing the municipal borrowing costs of uninsured bonds and leave insured bonds out of their samples. If researchers insist on including insured bonds in their sample with uninsured bonds, their analysis should take into consideration that default risk premiums are not the same for insured and uninsured bonds.

Today a large number of new issues are insured. So researchers may have difficulty finding a sufficiently large sample of uninsured bonds to perform an analysis of the interest cost of new bond issues. Researchers may have to resort to analyzing bond ratings since the underlying credit rating is published for

many governments who issue insured debt. As shown in Marks and Raman (1987), ordinal probit analysis is preferred to discriminant analysis in bond rating studies since the coefficients of the variables in the probit model can be tested for statistical significance. Researchers who have access to seasoned bond data should focus only on those bonds that are not insured. Marks et al. (1994) and Marks and Raman (1998) have shown that abnormal returns calculated from seasoned bonds can be used to examine the impact of specific events on municipal bond prices. They looked at two different types of events, bond rating changes and the California Tax Revolt. Ingram and Wilson (1999) investigated the relationship between municipal bond yields calculated from seasoned bond prices and accounting variables. They concluded that bond ratings and seasoned bond prices reflect accounting information primarily at the time new issues are sold.

REFERENCES

Ambac Financial Group, Inc., 10-K Report (1997).

Benson, E. D. (1979). The search for information by underwriters and its impact on municipal interest cost. *Journal of Finance*, *34*(September), 871–885.

Benson, E. D. (1999). The continuing cost to municipalities of selling competitive tax-exempt bonds by NIC. *Municipal Finance Journal*, *20*(Spring), 35–67.

Cook, T. Q. (1982). Determinants of individual tax-exempt bond yields: A survey of the evidence. *Economic Review, Federal Reserve Bank of Richmond*, *68*(May/June), 14–39.

Copeland, R. C., & Ingram, R. W. (1982). The association between municipal accounting information and bond rating changes. *Journal of Accounting Research*, *20*(Autumn), 275–289.

Fabozzi, F. J. (2000). *Bond markets, analysis and strategies* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Gunst, R. F., & Mason, R. L. (1980). *Regression analysis and its application: A data oriented approach*. New York: Marcel Dekker.

Hillman, A. M. (1936). *Municipal bonds: A century of experience*. New York: Prentice-Hall.

Ingram, R. W., & Wilson, E. R. (1999). Information intermediation and seasoned municipal bond yields. *Research in Government and Nonprofit Accounting*, *10*, 1–31.

Marks, B. R., & Raman, K. K. (1987). The information content of unfunded pension obligations for municipal bond ratings: An empirical evaluation. *Advances in Accounting*, *4*, 33–42.

Marks, B. R., & Raman, K. K. (1998). The application of event methodology in municipal finance research. *Municipal Finance Journal*, *19*(Spring), 1–15.

Marks, B. R., Raman, K. K., & Wilson, E. R. (1994). The effect of municipal bond rating change announcements on seasoned bond prices. *Municipal Finance Journal*, *15*(Fall), 17–35.

Raman, K. K. (1981). Financial reporting and municipal bond rating changes. *The Accounting Review*, *56*(October), 910–926.

Smith, A. T. (2001). Bond insurance covers nearly 50% of 1st-half new debt. *Bond Buyer*, *337*(August 13), 6.

U.S. Census Bureau (2000a). *Finances of municipal and township governments 1997*. U.S. Department of Commerce.

U.S. Census Bureau (2000b). *Public education finances 1997*. U.S. Department of Commerce.

Wallace, W. A. (1981). The association between municipal market measures and selected financial reporting practices. *Journal of Accounting Research*, *19*(Autumn), 502–520.

White, H. A. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometric*, *48(*May), 817–838.

Wilson, E. R., & Howard, T. P. (1984). The association between municipal market measures and selected financial reporting practices: Additional evidence. *Journal of Accounting Research*, *22*(Spring), 207–224.

UNDERWRITER PRESTIGE, STATE TAXES, AND REOFFERING YIELDS OF MUNICIPAL BOND OFFERINGS

Arthur Allen and George Sanders

ABSTRACT

We examine the relationship between underwriter prestige and reoffer yields of new issue local government bonds using a sample of 2,072 issues from 1999/2000. Underwriter prestige is measured as the rank of the size of total issues handled by the top 100 underwriters. We find that underwriter prestige is associated with higher reoffer yields. We provide evidence consistent with one explanation of this finding; the tendency of smaller underwriters to underwrite bonds that are tax-advantaged to local investors.

1. INTRODUCTION

A number of studies have examined the role of securities underwriters in the pricing of new issues. The theoretical arguments are based on signaling theory: The reputation of the underwriter is hypothesized to provide information to the market about otherwise unobservable quality characteristics of the underwritten securities. Because the municipal securities markets are characterized by infrequent trading of seasoned securities and infrequent releases of information by securities issuers, particularly for small issuers (Cook, 1982), signaling could play a more important role in that market. Because of the paucity of information releases by issuers,

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 51–70

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11003-3

information intermediaries may be able to provide signals to reduce information asymmetry between buyer and seller.

Carter and Manaster (1990), found evidence that the corporate securities markets pay a premium for IPOs managed by more prestigious underwriting firms. However, a study by Roden and Bassler (1996) found no such effect for new issues of municipal bonds. Roden and Bassler suggest that their results are consistent with segmentation of the market for municipal bonds (e.g. Feroz & Wilson, 1992), with issuers having monopoly power and "Issues with local interest dominate the market," (Roden & Bassler, 1996, p. 661) so that larger, more prestigious underwriters may not be able to extract a benefit from their reputational capital. However, they offer no explanation for the segmentation.

This paper extends Roden and Bassler (1996) by using a larger sample (2,072 compared to 409), more recent data (1999–2000 compared to the period 1977 to mid-1982), complete data for each observation (Roden & Bassler included observations with missing data), a wider range of entities, and an alternative measure of underwriter prestige. Unlike Roden and Bassler and contrary to signaling theory, we find that more prestigious underwriters sell bonds at higher interest rates. To explain this seemingly anomalous result, we examine Roden and Bassler's conclusion that the municipal bond market is segmented geographically.

One important reason for geographic segmentation in the local government bond market is the variation in taxability of interest. When states do not tax interest on in-state local government bond issues but do tax out-of-state local government bond issues, local investors are expected to have a strong preference for in-state bond issues. We refer to these states as tax-preferenced states. Several studies have found that bonds in tax-preferenced states sell at lower yields than bonds in other states (e.g. Lovely & Wasylenko, 1992).

We split our sample into tax-preferenced states vs. other states. Our results for tax-preferenced states are consistent with our full sample; high prestige underwriters sell bonds at higher yields. However, our sample in other states show that there is no difference in yields for high prestige underwriters. Overall, our results confirm Roden and Bassler's finding that prestigious underwriters of local government bonds are not associated with lower yields and, therefore, do not signal the quality of issues in the same manner as found by Carter and Manaster (1990) in the corporate equities market. Instead, we interpret these results as indicating that prestigious underwriters tend to underwrite bonds sold to out-of-state investors and, therefore, yields are often higher than when local underwriters sell bonds to local investors in tax-advantaged states.

The remainder of this paper is organized as follows. Section 2 discusses underwriter reputation and signaling theory. Section 3 discusses geographic segmentation in the local government bond market. Section 4 discusses the

methodology. Section 5 discusses the results, and the final section provides a summary and conclusions.

2. SIGNALING AND UNDERWRITER REPUTATION

If the quality of units of a good cannot be observed by buyers at the time of purchase, but the quality is revealed after purchase, the goods will trade in the market at a price reflecting the average quality. As Akerlof (1970) showed, producers have an incentive in such a market to reduce quality if in doing so they can reduce cost; sellers of above average quality goods who have high opportunity costs will withdraw from the market (adverse selection) setting in motion a process that will result in a lower quality of product being offered and that may result in only the lowest quality producers remaining in the market. However, Spence (1973) showed that if high quality producers can engage in an activity that is less costly for them than for low quality producers and is observable by the buyer before the purchase, then this activity can serve as a signal revealing a higher quality good which commands a higher price. Such costly signaling has been proposed as a explanation for a variety of behaviors, including the selection of a firm's capital structure and dividend policy (Ross, 1977).

Other research has examined the signaling role of third party certification. Third party certification of quality may be effective if there are no direct signals that can be provided by a producer. In this case a credible signal can be provided by an informed third party that puts its reputational capital at risk. Megginson and Weiss (1991) summarize three conditions for effective third party certification: (1) the third party must put reputational capital at risk if it certifies a producer as high quality when the producer is of a low quality; (2) the loss in value of reputational capital in the event of a false certification must exceed the value that could be obtained by the certifier from a false certification; (3) the producer's cost of obtaining certification must be such that a separating equilibrium is possible between low and high quality producers. Third party certification is the basis of research on the selection of auditors (Copley et al., 1995; Palmrose, 1986; Simunic, 1980; Titman & Trueman, 1986), and the selection of underwriters (Carter & Manaster, 1990).

Carter and Manaster (1990), drawing on Rock (1986), Beatty and Ritter (1986) and Titman and Trueman (1986), argue that prestigious underwriters are associated with initial public offerings (IPOs) with a lower dispersion of firm values and, hence, risk. They expect that underwriter prestige is positively associated with the marketing of low risk IPOs and that low risk issuers will receive a smaller IPO discount with less of a post issue run-up. Empirically, Carter and Manaster (1990)

measured underwriter reputation using a ranking based on the relative position of an underwriter's name in the tombstone announcement of IPOs. They found that both price run-up and the standard deviation of price run-up is individually significantly negatively related to their measure of underwriter prestige. They interpret the results as consistent with a more accurate pricing of the securities.

Roden and Bassler (1996) examined the effect of underwriter prestige on the interest cost of municipal bond offerings in 409 municipal offerings. They used the Carter and Manaster (1990) technique of measuring underwriter prestige by an underwriter's relative position in a tombstone ad. However, their results were inconsistent with the previous literature in the IPO market; they did not find a significant effect on bond yields associated with a city's use of a prestigious underwriter. This is inconsistent with the findings of Carter and Manaster (1990). We use a more readily available measure of underwriter prestige based on the volume of issues underwritten during the sample period (1999–2000). An advantage of using underwriter size as a proxy is that it is an objectively observable, interval scale variable, while the expert opinion rankings and positions in tombstone advertisements used in CM require the use of relatively arbitrary judgments. Also, we use a measure of the price paid by the buyer (reoffer yields) rather than a measure of the cost to the issuer (NIC). Therefore, our methods provide a more direct test of signaling theory in the municipal bond market.

There is reason to believe that firm size may proxy for reputation. An idea well developed in the audit market literature (e.g. Simunic, 1980) is that larger firms have a larger investment in reputational capital on which they are able to earn rents. In this market, like the market for local debt securities, quality is hard to observe. Audited companies signal quality by selecting an auditor with a given reputation. Larger audit firms specialize in providing a higher quality, and size is often used in empirical studies as a quality proxy. A body of empirical work has demonstrated that larger audit firms appear able to charge higher fees, interpreted as a return on the reputational capital (e.g. Copley et al., 1995; Palmrose, 1986; Rubin, 1988; Simunic, 1980).

3. SEGMENTATION IN THE TAX-EXEMPT BOND MARKET

The municipal bond market is widely believed to be geographically segmented. For example, Mysak (1998) remarks that:

> Most of the market is state specific. In other words, a Florida municipal bond firm might spend all its time dealing strictly with the issuers and buyers in its own state. It is not too far off

> the market to say that, because of their independent, decentralized nature, nearly every state has a highly developed municipal bond market all its own (p. 21). . . . If you are like most issuers (i.e. small), you have a very local constituency . . . But say you are a larger issuer . . . You become national, in what we have seen is a largely regional market. . . . How will your bonds be distributed nationally? Again, using the Massachusetts issue as an example, fully 40% of the deal went to New York, 20% to Illinois, 13% to California, and 5% to Pennsylvania. Approximately 20% of the issue remained in Massachusetts. This too seems typical for a fairly large bond issue (pp. 127–128).

A number of studies have provided evidence that the municipal securities market is segmented geographically (Brucato et al., 1991; Hendershott & Kidwell, 1978; Kidwell et al., 1987; Leonard, 1983). These authors found evidence that in-state supply factors, bank pledging requirements and state income tax rates and policies were associated with differences in local governments' securities' yields. When states do not tax interest on in-state local government bond issues but do tax out-of-state local government bond issues, local investors are expected to have a strong preference for in-state bond issues. We refer to these states as tax-preferenced states.

We expect that more prestigious underwriters will have a larger, more geographically diverse sales force than other underwriters. Larger underwriters are also expected to have developed closer relationships with institutional buyers. These institutional buyers tend to be geographically concentrated into the larger financial markets which exist in only a small number of states. Because of a limited supply of local government bonds in each state and the desire of institutional investors to diversify, institutional buyers often purchase bonds from out-of-state issuers. All of the above factors suggest that the larger, more prestigious underwriters will have a comparative advantage at underwriting bonds that are likely to be sold out-of-state.

Because prestigious underwriters have a comparative advantage at selling bonds to out-of-state investors and bonds in tax-preferenced states tend to be sold to local investors at lower interest rates, it is important to control for state tax policies when investigating the effect of underwriter prestige in the local government bond market. If underwriter prestige signals the quality of the issue, we believe this effect will be strongest in the non-tax-preferenced states. Therefore, we split our sample into tax-preferenced vs. other states. Regression models are run for the overall sample and for each subsample.

4. METHODOLOGY

We use a regression model based on earlier work in the analysis of the municipal bond markets (e.g. Leonard, 1999; Roden & Bassler, 1996). The dependent variable is the reoffer yield of the new issue. The independent variables represent

issuer characteristics, market characteristics, underwriter prestige and state tax policies. Our data was drawn from the Bond Buyer Online over a period beginning January 20, 1999 and ending December 31, 2000. We downloaded sale information for all competitive and negotiated general obligation bond issues. Revenue bonds were excluded because roeffer yields can vary widely depending on issue and issuer characteristics that cannot be captured by the available data. Bond ratings were obtained from Moody's Bond Record. A complete data set was obtained for 2,072 issues.

4.1. Reoffer Yields

We used the 10 year reoffering yield to measure the interest cost of the issue.[1] We choose to use reoffering yield instead of TIC (true interest cost) because the reoffering yield most directly measures the behavior of bond buyers. Also, Simonsen, Robbins, and Jump (2001) provide several criticisms of TIC including the fact that underwriting costs are sometimes included in the calculation of TIC while at other times they are excluded.

4.2. Underwriter Prestige

Roden and Bassler (1996) tested two reputation index numbers based on tombstone advertisements and an index based on ratings by an expert panel. Interestingly, the two tombstone index numbers were inversely correlated with the expert panel. Roden and Bassler (1996) suggest that the panel of experts may have been confused by reputation in the common stock new issues market and they chose not to use the expert opinions in further analysis. Instead of tombstone advertisements, we use underwriter size as our measure of underwriter prestige. We measure underwriter size as the underwriter's rank in the dollar amount underwritten of long-term local government bonds during 1999/2000.[2] The top 100 underwriters accounted for 213 out of 219 billion dollars of new local government bonds issued during 1999. The few underwriters not in the top 100 were coded 101. Our measure of underwriter prestige is inverse; the higher the prestige, the lower the underwriter rank.

4.3. State Taxation of Local Government Interest Payments

State taxation of local government interest payments varies widely. Some states have no income taxation or impose no taxes on any local government interest

payments. Other states impose taxes on all interest payments. Previous research (e.g. Lovely & Wasylenko, 1992) in this area groups states into tax-preferenced vs. other. Tax-preferenced states are those states that tax local government interest payments on bonds issued out-of-state but not in-state. We report separate models for tax-preferenced vs. other states. Because tax policies can change over time, we requested personal income tax forms from all 50 states rather than relying on the classifications used by prior research.

4.4. Issuer and Market Characteristics

As discussed previously, we include a dummy variable in the regression to control for competitive vs. negotiated issues. Benson (1979) analyzed a sample of municipal issues and found evidence that interest cost varies inversely with the number of bidders for competitive issues. We control for this underwriter competition effect by including the log of the number of bidders listed in the online Bond Buyer for each issue. The number of bidders is set to 1 for negotiated issues.

To control for the level of interest rates for local government bonds, we used the Bond Buyer's Municipal Bond Index yield to maturity. Yield curves of municipal bonds are monotonically upward sloping. We use the log of maturity in months to control for the yield curve. Call provisions reduce the expected duration of the bond and therefore are expected to increase yields. We include a variable coded 1 for callable bonds and 0 otherwise.

To control for default risk, dummy variables are included for Moody's Aa, A, Baa, and Ba ratings with Aaa ratings the excluded group. A dummy variable is also included for bonds not rated by Moody's because they are expected to sell for higher yields (Reeve & Herring, 1986). Although insured bonds generally receive a Aaa rating, they tend to trade at rates higher than uninsured Aaa bonds (e.g. Hsueh & Chandy, 1989). Therefore, we include a dummy variable coded 1 for insured bonds and 0 for other bonds.

Although we do not include revenue bonds, we do include bonds from counties and school districts as well as municipalities. We include separate dummy variables for counties and school districts with municipalities the excluded group. Refunding bonds are typically associated with less risk than other issues so we include a dummy variable coded 1 for refunding bonds and 0 for other issues. Bonds that are bank qualified are in high demand and therefore are expected to sell for lower yields. We code a dummy variable 1 for bank qualified bonds and 0 for other bonds.

We also include several demographic variables. Because of limited data availability, we present results with a full sample without these variables and a

subsample where these data are available. Housing units per capita measures availability of housing. Vacancy rate may proxy for growth, infrastructure available for expansion, economic vibrancy, etc. Percentage of Population with Bachelor's Degree proxies for population educational attainment. Affordable housing, median housing values, and median household income represents both the tax base and the ability of taxpayers to afford taxes. Population could proxy for several factors including the sophistication of the issuer and the market's familiarity with the issuer. Percentage unemployed measures both the demand for services and the ability of taxpayers to pay.

5. RESULTS

5.1. Full Sample

Descriptive statistics are presented in Table 1. The full sample includes 2,072 issues. If there is more than one issue per issuer, only the most recent issue is retained in the sample. The sample size is significantly reduced for the demographic variables ($n = 644$) in part because we excluded school districts. The data for school districts was generally unavailable, and the effect of the demographic variables is expected to be very different for school districts vs. cities.

The regression results for the full sample are reported in Table 2. The adjusted *R*-square is quite high (0.88) and diagnostic tests indicate no violations of assumptions underlying OLS regression. All variables are significant in the expected direction except that County is not significant, Bank Qualifed is marginally significant (p-value $= 0.153$ for a two-tailed test) and the Prior Issues variable is positive. The tax preference variable ($t = -7.82$) indicates that yields are lower when bonds are issued in tax-preferenced states. The underwriter variable is significantly negative ($t = -2.4$). Because more prestigious underwriters are coded lower than less prestigious underwriters, a negative sign on this variable implies that underwriter prestige is associated with higher yields. This result is counter to the idea that more prestigious underwriters signal higher quality issues. To investigate this seemingly anomalous result, we split the sample into tax-preferenced vs. other states.

Table 3 reports the regression results for the non-tax-preferenced states. The tax variable is necessarily deleted from this analysis. Ratings variables for Baa and Ba were omitted because there were no observations with these ratings for this group. The adjusted *R*-square remains high (0.88). There are 547 observations for this group. The underwriter variable is no longer significant ($t = -0.28$). This result is consistent with Roden and Bassler (1996) who found that underwriter

Table 1. Descriptive Statistics for a Sample of Negotiated and Competitively Bid General Obligation Bonds Issued by Local Governments in 1999 and 2000.

Variable	Description	Mean	Std. Dev.	Minimum	Maximum
Panel A: Variables related to the issuer and issues ($n = 2{,}072$)					
Reoffer yield	10-year reoffer yield (reoffering yield closest to 10 years if the 10 year reoffer yield is unavailable)	4.85	0.38	3.45	6.05
Interest rate	Bond Buyer weekly index of long-term municipal bond interest rates	5.56	0.31	4.96	6.11
Log of maturity	Natural log of the maturity (in months) of the associated reoffer yield	4.76	0.17	3.18	5.70
Log of bids	The natural logarithm of the number of bids for competitive issues (coded 0 for negotiated issues)	1.34	0.55	0.69	3.14
Callable	Coded 1 if the issue is callable	0.91	0.29	0.00	1.00
Bank qualified	Coded 1 if the issue is bank qualitifed	0.42	0.49	0.00	1.00
Insurance	Coded 1 if the issue is insured	0.49	0.50	0.00	1.00
PSF insurance	School district bonds guaranteed by the state of Texas	0.03	0.17	0.00	1.00
Aa rating	Coded 1 if the issue rated Aa	0.16	0.37	0.00	1.00
A rating	Coded 1 if the issue rated A	0.10	0.30	0.00	1.00
Baa rating	Coded 1 if the issue rated Baa	0.02	0.14	0.00	1.00
Ba rating	Coded 1 if rated Ba	0.00	0.05	0.00	1.00
Unrated	Coded 1 if the issue is unrated	0.18	0.38	0.00	1.00
Refunding bond	Coded 1 if the issue is a refunding issued	0.09	0.29	0.00	1.00
Competitive issue	Coded 1 if the issue is sold by competitive bid	0.84	0.37	0.00	1.00
County	Coded 1 if the issuer is a county	0.14	0.35	0.00	1.00
School district	Coded 1 if the issuer is a school district	0.31	0.46	0.00	1.00
Smallest issue	Coded 1 if the issue size is under $1,000,000	0.10	0.29	0.00	1.00
Smaller issue	Coded 1 if the issue size is between $1,000,000 and $2,500,000	0.19	0.39	0.00	1.00
Small issue	Coded 1 if the issue size is between $2,500,000 and $5,000,000	0.22	0.42	0.00	1.00

Table 1. (*Continued*)

Variable	Description	Mean	Std. Dev.	Minimum	Maximum
Large issue	Coded 1 if the issue size is between $10,000,000 and $25,000,000	0.15	0.35	0.00	1.00
Larger issue	Coded 1 if the issue size is over $25,000,000	0.11	0.32	0.00	1.00
Log of previous issues	Natural log of the number of prior issues by that issuer in the ten years prior to issuance	0.80	1.02	0.00	4.84
Tax preference	Coded 1 if issued in a state in which interest income is not taxed on in-state bond but is taxed on out-of-state bonds issues	0.74	0.44	0.00	1.00
Underwriter rank	Coded 1–101 for the rank of the underwriter by the total dollar amounts underwritten during the year of issuance for the top 100 underwriters with unranked underwriters coded 101	40.50	35.78	1	101
Panel B: Demographic variables ($n = 644$)					
Housing units per capita	Total housing units divided by total population, Bureau of the Census: Census 2000	0.41	0.06	0.07	0.75
Vacancy rate	Unoccupied housing units divided by the total units, Bureau of the Census: Census 2000	0.07	0.05	0.01	0.42
% of Pop. with Bachelor's degree	Bureau of the Census: Census 2000	0.11	0.05	0.00	0.28
Affordable housing	Median value of housing divided by median family income, 2000, Bureau of the Census: Census 2000	5.63	1.69	1.79	18.17
Population	Bureau of the Census: 2000 Decennial Profiles	256,071	817,698	141	8,008,278
% Unemployed	Bureau of the Census: 2000 Decennial Profiles	5.01	2.26	0	13.90
Median houshold income	Bureau of the Census: 2000 Decennial Profiles	47,211	16,751	18,218	136,462
Median housing value	Bureau of the Census: 2000 Decennial Profiles	135,641	76,678	29,600	662,400

Table 2. Regression Results for a Sample of 2,072 Negotiated and Competitively Bid General Obligation Bonds Issued by Local Governments in 1999 and 2000.

Variable	Parameter Estimate	Standard Error	*t*-Statistic	Probability
Intercept	−3.561	0.106	−33.520	0.000
Interest rate	1.004	0.011	92.390	0.000
Maturity	0.600	0.018	32.580	0.000
Log of bids	−0.038	0.006	−6.300	0.000
Call	0.034	0.011	3.200	0.001
Bank qualified	−0.012	0.008	−1.430	0.153
Insurance	0.068	0.015	4.470	0.000
PSF insurance	0.034	0.020	1.760	0.079
Aa rating	0.035	0.016	2.180	0.029
A rating	0.039	0.017	2.220	0.026
Baa rating	0.223	0.026	8.510	0.000
Ba rating	0.234	0.062	3.790	0.000
Unrated	0.102	0.017	5.870	0.000
Refunding	−0.032	0.011	−2.960	0.003
Competitive issue	−0.079	0.010	−8.190	0.000
County	−0.009	0.009	−1.050	0.292
School district	−0.076	0.008	−10.020	0.000
Smallest issue	0.237	0.014	17.180	0.000
Smaller issue	0.060	0.010	6.090	0.000
Small issue	0.024	0.009	2.690	0.007
Large issue	0.045	0.010	4.330	0.000
Larger issue	0.055	0.012	4.590	0.000
Prior issues	0.007	0.003	2.030	0.042
Tax preference	−0.057	0.007	−7.900	0.000
Underwriter	−0.022	0.009	−2.400	0.017

Note: The dependent variable is the reoffering yield of the issue. See Table 1 for a description of the variables.
$n = 2{,}072$. Adjusted R-square $= 0.88$.

prestige is not associated with interest cost for municipal bond issues. Our results confirm that underwriter prestige does not signal the quality of the issue for local government bonds.

Table 4 reports the regression results for the tax-preferenced states. There are 1,525 observations in this group, and the adjusted R-square remains high (0.88). As in the full sample, the underwriter prestige variable is significantly negative ($t = -2.31$), and the magnitude of the coefficient is quite similar to that in the full sample. Although we find a positive association between underwriter prestige and yields, we do not believe that more prestigious underwriters cause higher yields. Instead, we believe that because this result only holds in tax-preferenced

Table 3. Regression Results for a Sample of 547 Negotiated and Competitively Bid General Obligation Bonds Issued between 1999 and 2000 by Local Governments in States with No Tax Preference for In-State Bonds.

Variable	Parameter Estimate	Standard Error	*t*-Statistic	Probability
Intercept	−3.567	0.207	−17.260	0.001
Interest rate	1.018	0.021	48.340	0.001
Maturity	0.581	0.036	16.250	0.001
Log of bids	−0.047	0.011	−4.410	0.001
Call	0.006	0.029	0.210	0.833
Bank qualified	−0.023	0.016	−1.440	0.150
Insurance	0.102	0.033	3.070	0.002
PSF insurance	0.028	0.023	1.220	0.221
Aa rating	0.030	0.035	0.840	0.402
A rating	0.073	0.039	1.900	0.058
Unrated	0.118	0.037	3.170	0.002
Refunding	−0.012	0.020	−0.600	0.546
Competitive issue	−0.029	0.019	−1.470	0.143
County	−0.020	0.019	−1.030	0.305
School district	−0.090	0.016	−5.530	0.001
Smallest issue	0.205	0.029	7.180	0.001
Smaller issue	0.046	0.020	2.330	0.020
Small issue	0.041	0.018	2.310	0.022
Large issue	0.089	0.023	3.960	0.001
Larger issue	0.076	0.023	3.350	0.001
Prior issues	0.001	0.007	0.060	0.956
Underwriter	−0.006	0.020	−0.280	0.778

Note: The dependent variable is the reoffering yield of the issue. See Table 1 for a description of the variables.
$n = 547$. Adjusted R-square = 0.88.

states, the association is caused by the tendency of prestigious underwriters to underwrite bonds being sold primarily to out-of-state buyers. Consistent with the remarks by Mysak (1998), local underwriters are able to sell bonds at lower average yields in tax-preferenced states because they sell more bonds to in-state buyers.

5.2. Reduced Sample with Demographic Data

Table 5 reports the results for the subsample where demographic data are available. The sample size is significantly reduced ($n = 644$). However, the adjusted R-square is still high (0.85). Of the demographic variables, only the housing units per capita is significant at the 0.10 level. The results for the key

Table 4. Regression Results for a Sample of 1,525 Negotiated and Competitively Bid General Obligation Bonds Issued between 1999 and 2000 by Local Governments in States with a Tax Preference for In-State Bonds.

Variable	Parameter Estimate	Standard Error	*t*-Statistic	Probability
Intercept	−3.579	0.125	−28.640	0.001
Interest rate	1.000	0.013	78.020	0.001
Maturity	0.598	0.022	27.620	0.001
Log of bids	−0.034	0.007	−4.630	0.001
Call	0.039	0.012	3.370	0.001
Bank qualified	−0.010	0.010	−1.080	0.280
Insurance	0.058	0.017	3.370	0.001
Aa rating	0.036	0.018	2.030	0.043
A rating	0.030	0.020	1.510	0.131
Baa rating	0.213	0.027	7.800	0.001
Ba rating	0.222	0.062	3.590	0.001
Unrated	0.097	0.020	4.910	0.001
Refunding	−0.040	0.013	−3.050	0.002
Competitive issue	−0.094	0.011	−8.240	0.001
County	−0.006	0.010	−0.560	0.578
School district	−0.073	0.009	−8.500	0.001
Smallest issue	0.242	0.016	15.250	0.001
Smaller issue	0.065	0.011	5.720	0.001
Small issue	0.017	0.010	1.720	0.086
Large issue	0.035	0.012	2.930	0.004
Larger issue	0.048	0.014	3.380	0.001
Prior issues	0.008	0.004	1.940	0.053
Underwriter	−0.024	0.010	−2.310	0.021

Note: The dependent variable is the reoffering yield of the issue. See Table 1 for a description of the variables.
$n = 1{,}525$. Adjusted R-square $= 0.88$.

variables are very similar to the full sample. The tax preference variable remains significant and its magnitude is about the same. The underwriter variable also remains insignificant. Table 6 reports the results for the issues in states with no tax preference for in-state bonds and demographic data available. Like the Table 5 sample, the underwriter variable is insignificant. Table 7 reports the results for the issues in states with a tax preference for in-state bonds and demographic data available. Like the full sample, the underwriter variable is significantly negative (p-value $= 0.09$). The magnitude of this coefficient is larger in the full sample (-0.032 compared to -0.024), although the t-statistic is smaller because of the reduced sample size (1,525 comapred to 479). The results for the reduced sample using the demographic variables are consistent with those of the full sample.

Table 5. Regression Results for a Sample of 644 Negotiated and Competitively Bid General Obligation Bonds Issued between 1999 and 2000 by Local Governments: All Issues with Demographic Data Available.

Variable	Parameter Estimate	Standard Error	*t*-Statistic	Probability
Intercept	−3.404	0.225	−15.160	0.000
Interest rate	0.954	0.020	48.580	0.000
Maturity	0.610	0.037	16.680	0.000
Log of bids	−0.035	0.010	−3.570	0.000
Call	0.006	0.019	0.310	0.757
Bank qualified	−0.045	0.015	−3.040	0.003
Insurance	0.079	0.023	3.410	0.001
PSF insurance	0.164	0.133	1.230	0.218
Aa rating	0.032	0.023	1.420	0.156
A rating	0.049	0.026	1.870	0.063
Baa rating	0.345	0.053	6.480	0.000
Ba rating	0.173	0.082	2.110	0.035
Unrated	0.103	0.029	3.510	0.001
Refunding	−0.021	0.021	−0.960	0.339
Competitive issue	−0.091	0.019	−4.690	0.000
County	0.003	0.014	0.190	0.848
Smallest issue	0.205	0.026	7.840	0.000
Smaller issue	0.063	0.018	3.520	0.001
Small issue	0.030	0.016	1.840	0.067
Large issue	0.043	0.019	2.300	0.022
Larger issue	0.044	0.020	2.260	0.024
Prior issues	0.012	0.005	2.400	0.017
Tax preference	−0.056	0.013	−4.260	0.000
Housing units per capita	0.357	0.206	1.730	0.084
Percent vacant housing	−0.294	0.187	−1.570	0.117
Percent of population with Bachelor's degree	−0.336	0.229	−1.470	0.142
Affordability of housing	−0.022	0.018	−1.220	0.224
Population	0.000	0.000	1.490	0.138
Percent unemployed	−0.001	0.003	−0.360	0.718
Per capita income	0.000	0.000	0.030	0.979
Median value of housing	0.000	0.000	1.000	0.319
Underwriter	−0.017	0.017	−1.000	0.320

Note: The dependent variable is the reoffering yield of the issue. See Table 1 for a description of the variables.
$n = 644$. Adjusted R-square $= 0.85$.

5.3. Models by Type of Issuer

The effect of underwriter and taxes might vary depending on whether the issuer is a city, county, or school district. This section discusses results from subsamples for cities, counties, and school districts. Table 8 reports the results for the underwriter

Table 6. Regression Results for a Sample of 165 Negotiated and Competitively Bid General Obligation Bonds Issued between 1999 and 2000 by Local Governments in States with No Tax Preference for In-State Bonds: All Issues with Demographic Data Available.

Variable	Parameter Estimate	Standard Error	*t*-Statistic	Probability
Intercept	−4.278	0.502	−8.520	0.000
Interest rate	1.002	0.045	22.480	0.000
Maturity	0.645	0.062	10.440	0.000
Log of bids	−0.074	0.021	−3.560	0.001
Call	−0.021	0.050	−0.420	0.675
Bank qualified	−0.098	0.033	−3.000	0.003
Insurance	0.058	0.075	0.770	0.443
PSF insurance	0.151	0.152	0.990	0.322
Aa rating	0.017	0.073	0.230	0.816
A rating	0.044	0.081	0.540	0.591
Unrated	0.096	0.085	1.140	0.257
Refunding	−0.004	0.042	−0.100	0.920
Competitive issue	−0.062	0.046	−1.350	0.180
County	−0.014	0.032	−0.440	0.663
Smallest issue	0.156	0.061	2.560	0.012
Smaller issue	0.092	0.039	2.350	0.020
Small issue	0.094	0.035	2.670	0.009
Large issue	0.152	0.047	3.220	0.002
Larger issue	0.020	0.052	0.390	0.699
Prior issues	−0.012	0.012	−0.990	0.325
Underwriter	0.024	0.037	0.650	0.517
Housing units per capita	1.129	0.544	2.080	0.040
Percent vacant housing	−0.878	0.528	−1.660	0.099
Percent of pop. with Bachelor's degree	−1.002	0.481	−2.090	0.039
Affordability of housing	0.013	0.016	0.780	0.436
Population	0.000	0.000	0.410	0.680
Percent unemployed	0.008	0.008	1.030	0.304
Per capita income	0.000	0.000	2.100	0.037
Median value of housing	0.000	0.000	−0.830	0.409

Note: The dependent variable is the reoffering yield of the issue. See Table 1 for a description of the variables.
$n = 165$. Adjusted R-square $= 0.86$.

variable for the full sample and for restricted samples using only cities, counties, or school districts.

Like the full sample, the underwriter variable for these government-type subsamples is not significant for issues in states with no tax preference. However, in states with a tax preference in both the city subsample and in the school

Table 7. Regression Results for a Sample of 479 Negotiated and Competitively Bid General Obligation Bonds Issued between 1999 and 2000 by Local Governments in States with a Tax Preference for In-State Bonds: All Issues with Demographic Data Available.

Variable	Parameter Estimate	Standard Error	*t*-Statistic	Probability
Intercept	−3.108	0.323	−9.630	0.000
Interest rate	0.951	0.023	42.190	0.000
Maturity	0.568	0.049	11.690	0.000
Log of bids	−0.023	0.012	−1.950	0.052
Call	0.016	0.021	0.740	0.461
Bank qualified	−0.039	0.017	−2.240	0.026
Insurance	0.081	0.025	3.240	0.001
Aa rating	0.039	0.024	1.600	0.110
A rating	0.037	0.028	1.310	0.190
Baa rating	0.339	0.054	6.340	0.000
Ba rating	0.160	0.082	1.950	0.051
Unrated	0.250	0.026	9.750	0.000
Refunding	0.103	0.033	3.130	0.002
Competitive issue	−0.039	0.025	−1.520	0.129
County	−0.105	0.022	−4.710	0.000
Smallest issue	−0.001	0.016	−0.050	0.963
Smaller issue	0.219	0.029	7.440	0.000
Small issue	0.052	0.020	2.550	0.011
Large issue	0.017	0.018	0.930	0.350
Larger issue	0.014	0.021	0.660	0.507
Prior issues	0.038	0.022	1.700	0.089
Underwriter	0.017	0.006	2.850	0.005
Housing units per capita	−0.033	0.019	−1.750	0.080
Percent vacant housing	0.052	0.308	0.170	0.865
Percent of population with Bachelor's degree	−0.053	0.236	−0.230	0.822
Affordability of housing	−0.309	0.238	−1.300	0.195
Population	−0.012	0.009	−1.380	0.169
Percent unemployed	0.000	0.000	1.110	0.267
Per capita income	−0.004	0.005	−0.820	0.411
Median value of housing	0.000	0.000	−0.090	0.931

Note: The dependent variable is the reoffering yield of the issue. See Table 1 for a description of the variables.
$n = 479$. Adjusted R-square $= 0.86$.

district subsample the coefficients for the underwriter variable are marginally significant (p-value $= 0.065$ and 0.093 for a one-tailed test). The magnitude of the underwriter variable is similar for cities, school districts and the full sample (−0.019, −0.029, −0.024, respectively).

Table 8. OLS Parameter Estimates (*t*-Statistics) for Underwriter by State Tax Preference and Government Type: Negotiated and Competitively Bid General Obligation Bonds Issued between 1999 and 2000 by Local Governments.

	Full Sample	City Governments	County Governments	School Districts
No tax preference states	−0.006 (−0.280) $n = 547$	−0.007 (−0.280) $n = 314$	0.003 (0.040) $n = 61$	−0.031 (−0.750) $n = 172$
Tax preference states	−0.024 (−2.310)** $n = 1{,}525$	−0.019 (−1.520) $n = 836$	−0.023 (−0.750) $n = 229$	−0.029 (−1.330) $n = 460$

**Significant at <0.05.

Unlike the full sample, the underwriter variable is not significant for counties in states with no tax preference (p-value = 0.227 for a one-tailed test), although the magnitude of the underwriter variable is similar for the county model and the full sample (−0.023 vs. −0.022). The samples sizes of the counties is quite small and, therefore, care should be taken in the interpretation of these results.

Overall, the results by issuer type are consistent with, but not as strong as, the full sample results. The underwriter variable is never significant using issues in states with no tax preference for in-state bonds. The underwriter variable is marginally significant for cities and school districts but not counties, and the magnitude of the underwriter variable's coefficient is similar for the full sample vs. each of the three issuer types (−0.022 full sample vs. −0.019, −0.023, and −0.029 for cities, counties, and school districts).

6. SUMMARY AND CONCLUSIONS

Signaling theory suggests that underwriter prestige signals the quality of a security issue. Therefore, underwriter prestige is expected to increase the price (decrease the yields) paid by the buyer. Empirical support for this relationship was found by Carter and Manaster (1990) in the corporate IPO market but not by Roden and Bassler (1996) in the municipal bond market. Roden and Bassler suggested that their results were driven by segentation in the local government bond market. This paper extends Roden and Bassler by using a larger sample (1,439 compared to 409), more recent data (1999 compared to the period 1977 to mid 1982), complete data for each observation (Roden & Bassler included observations with missing data), a wider range of entities, and an alternative measure of underwriter prestige. More importantly, we investigate the idea that the relationship between underwriter prestige and local government reoffer yields is driven by segmentation in the bond market.

We find that underwriter prestige is not associated with higher prices paid for local government bonds. To the contrary, we find that underwriter prestige is associated with higher reoffer yields. To examine whether this result is driven by segmentation of the market, we split the sample into tax-preferenced issues (from states that do not tax income on in-state bonds but do tax the income on out-of-state bonds) vs. other issues. We find that the unexpected association between underwriter prestige and reoffer yields only exists in tax-preferenced states. There is no association between underwriter prestige and reoffer yields in other states. This result provides some indirect evidence that larger underwriters may not compete as actively for issues in states where there is a tax preference for in-state bonds.

National underwriters have a nation-wide sales distribution network and, therefore, are expected to be most interested in issues that sell nation-wide. Issues made in states with a tax preference are believed to sell a larger portion of the issue locally because these bonds yield higher after-tax returns to local investors. Since the smaller underwriters have a more geographically limited sales force, smaller underwriters are expected to have a comparative advantage for bonds that are likely to be sold to local investors. Our results confirm Roden and Bassler's conclusions that issuers do not lower their interest costs by using a prestigious underwriter, and that "Issues with local interest dominate the market" (p. 660). The results are also consistent with the anecdotal observations of Mysak (1998). We find that the segmentation observed by Roden and Bassler may be a result of the differential tax treatment of municipal bond interest.

NOTES

1. We choose 10 years because this was the average maturity of the competitive issues and it was commonly available. Where a 10 year reoffering yield was not available, we used the reoffering yield closest to 10 years.
2. Our source was the Spring 2000 and 2001 editions of "The Bond Buyer's Municipal Marketplace." The Bond Buyer cites the source of this data as Thompson Financial Securities Data. To make the magnitude of the underwriter variable comparable to other variables, it is multiplied by 0.01.

ACKNOWLEDGMENTS

Research assistance by Wen Guo and Xuebing Tang is appreciated. Helpful suggestions from the editor and an anonymous reviewer are gratefully acknowledged.

REFERENCES

Akerlof, G. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*, 488–500.

Beatty, R. P., & Ritter, J. R. (1986). Investment banking, reputation, and the underpricing of initial public offerings. *Journal of Financial Economics*, *15*, 213–232.

Benson, E. D. (1979). The search for information by underwriters and its impact on municipal interest costs. *The Journal of Finance*, *34*, 871–885.

The Bond Buyer (2000–2001). *The bond buyer's municipal marketplace* (Spring). New York: American Banker.

The Bond Buyer (2000–2001). *The bond buyer online*. New York: American Banker. <http://www.bondbuyer.com/index.html>.

Brucato, P. F., Forbes, R. W., & Leonard, P. A. (1991). The effects of state tax differentials on municipal bond yields. *Municipal Finance Journal*, *12*, 59–77.

Bureau of the Census (2002). 2000 Decennial Profiles. May–June 2002. http://ftp2.census.gov/census_2000/datasets/100_and_sample_profile/ accessed August 20, 2002.

Carter, R., & Manaster, S. (1990). Initial public offerings and underwriter reputation. *The Journal of Finance*, *45*, 1045–1067.

Cook, T. Q. (1982). Determinants of individual tax-exempt bond yields: A survey of the evidence. *Economic Review Federal Reserve Bank of Richmond*, *68*, 14–40.

Copley, P. A., Gaver, J. J., & Gaver, K. M. (1995). Simultaneous estimation of the supply and demand of differentiated audits: Evidence from the municipal audit market. *Journal of Accounting Research*, 137–156.

Feroz, E. H., & Wilson, E. R. (1992). Market segmentation and the association between municipal financial disclosure and net interest cost. *The Accounting Review*, 480–495.

Hendershott, P. H., & Kidwell, D. S. (1978). The impact of relative security supplies: A test with data from a regional tax-exempt bond market. *Journal of Money, Credit and Banking*, *10*, 337–347.

Hsueh, L. P., & Chandy, P. R. (1989). An examination of the yield spread between insured and uninsured debt. *The Journal of Financial Research*, *12*, 235–243.

Kidwell, D. S., Koch, T. M., & Stock, D. R. (1987). Issue size and term structure segmentation effects on regional yield differentials in the municipal bond market. *Journal of Economics and Business*, *39*, 339–347.

Leonard, P. (1983). Some factors determining municipal revenue bond interest costs. *Journal of Economics and Business*, *35*, 71–83.

Leonard, P. A. (1999). Competitive bidding for municipal bonds: New tests of the underwriter search hypothesis. *Municipal Finance Journal*, *19*, 18–38.

Lovely, M. E., & Wasylenko, M. J. (1992). State taxation of interest income and municipal borrowing costs. *National Tax Journal*, *45*, 37–53.

Megginson, W. L., & Weiss, K. A. (1991). Venture capital certification in initial public offerings. *The Journal of Finance*, *46*, 879–903.

Mysak, J. (1998). *Handbook for muni bond issuers*. Princeton, NJ: Bloomberg Press.

Palmrose, Z.-V. (1986). Auditor fees and auditor size: Further evidence. *Journal of Accounting Research*, *24*, 97–110.

Reeve, J. M., & Herring, H. (1986). An examination of nonrated municipal bonds. *Journal of Economics and Business*, *38*, 65–76.

Rock, K. (1986). Why new issues are underpriced. *Journal of Financial Economics*, *15*, 187–212.

Roden, P. F., & Bassler, J. (1996). Effect of underwriter prestige on the interest cost of municipal bond offerings. *The Financial Review*, *31*, 641–666.

Ross, S. A. (1977). The determination of financial structure: The incentive signaling approach. *Bell Journal of Economics*, 23–40.

Rubin, M. A. (1988). Municipal audit fee determinants. *The Accounting Review*, *63*, 219–236.

Simonsen, B., Robbins, M. D., & Jump, B., Jr. (2001). Is TIC always the most appropriate cost of capital measure? *Municipal Finance Journal*, *22*, 1–13.

Simunic, D. A. (1980). The pricing of audit services: Theory and evidence. *Journal of Accounting Research*, *18*, 161–190.

Spence, A. M. (1973). *Market signaling: Information transfer in hiring and related screening processes*. Cambridge, MA: Harvard University Press.

Titman, S., & Trueman, B. (1986). Information quality and the valuation of new issues. *Journal of Accounting and Economics*, *16*, 3–50.

EFFECTS OF NONPROFIT ORGANIZATION WEALTH AND EFFICIENCY ON PRIVATE DONATIONS TO LARGE NONPROFIT ORGANIZATIONS

Nicholas P. Marudas

ABSTRACT

I examine the effects of wealth and efficiency of large nonprofit organizations (NPOs) on private donations to such NPOs by testing a model of private donations in pooled cross-sectional form and error components form (one-way fixed effects) using a panel of 1,239 NPOs divided into six industry samples. Results from testing the pooled cross-sectional model suggest, consistent with theory, that private donations of large arts and philanthropic NPOs are negatively related to NPO wealth. The results also suggest that private donations to large education NPOs are positively related to NPO wealth and to NPO efficiency. Results for other industry samples are considered to be inconclusive because they are not robust.

I use a one-way fixed effects specification to control for unspecified NPO-specific factors that may be correlated with the independent variables in the model, but find evidence of substantial measurement error in many of the variables in the model. This implies that obtaining valid results from

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 71–91

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11004-5

testing error components (fixed and random effects) models using this panel database, recently developed by the National Center for Charitable Statistics, is unlikely.

1. INTRODUCTION

Nonprofit organizations (NPOs) comprise a significant part of the U.S. economy. In 2000, total private donations to NPOs reached $230 billion, or 2.0% of GDP (AAFRC Trust for Philanthropy, 2001). Nearly all NPOs receive private donations, and for many NPOs, private donations are a significant source of revenue. Thus, understanding donors' giving decisions and the information used in these decisions is important. Relatively little research examines the impact of information calculated from NPO financial disclosures on private donations.

This study advances the literature on the effects that information calculated from financial disclosures of U.S. NPOs has on private donations to U.S. NPOs by:

(1) Developing theory for why an NPO's wealth may affect its private donations, and providing evidence on the effect of the financial ratio "years of available assets," a proxy for NPO wealth, on private donations of NPOs with total assets of at least $10 million (large NPOs). "Years of available assets" is "net assets"/("total expenses" – "fundraising expenses"), all of which are line items on NPO financial statements and IRS Forms 990.
(2) Providing evidence on the effect of other information, including the ratio "administrative efficiency" calculated from NPO financial disclosures, on private donations to large U.S. NPOs using a specification of administrative efficiency consistent with theory.
(3) Investigating the feasibility of providing improved evidence on the effects of other information, including the ratio "administrative efficiency" calculated from NPO financial disclosures, on private donations to large U.S. NPOs by testing a panel data set of large U.S. NPOs and an error components model (one-way fixed effects). Error components models reduce the effect of correlated omitted variables and collinearity among explanatory variables (Hsiao, 1986).

The panel database used in this study was recently developed by the National Center for Charitable Statistics (NCCS). This database is "expected to be the tax-exempt sector's equivalent to the Compustat databases . . ." (Gordon et al., 1999, p. 113). However, using the method of Griliches and Hausman (1986), I find evidence of large measurement error in most of the independent variables, biasing parameter estimates toward zero and significance tests against significance. Furthermore,

large measurement error in a few independent variables contaminates results for the other variables in the model, even those for which parameter estimates are significant (Greene, 1993). This suggests that the NCCS database may be limited in its usefulness.

Results from testing a pooled cross-sectional model, which does not suffer as much bias from measurement error as error components models, suggest that NPO wealth has the predicted significant negative effect on private donations to large arts and philanthropic NPOs. These results support managers and directors of large NPOs in these industries who believe that an NPO's wealth negatively impacts private donations. However, NPO wealth is *positively* related to private donations to large education NPOs. One explanation for this is donations in the form of endowments are positively related to wealth because of the way NPOs account for endowment donations, and this positive relation is greater for large education NPOs than for other types of large NPOs. This explanation is discussed more fully later. The effects of wealth on private donations to large NPOs in the other industries tested (health, human services, and "other") are inconclusive because results are not robust to a different specification of years of available assets, to different levels of censoring, or to rank regression. Results suggest that age has a significant positive effect on private donations to large education, health, human services, and "other" large NPOs. The effect of age on donations to large arts and large philanthropic NPOs is inconclusive because results are not robust. Results suggest that efficiency has the expected positive effect on private donations to education NPOs. The effect of efficiency on private donations to large NPOs in the other five industries tested is inconclusive because results are not robust.

2. RELATED LITERATURE

Only three studies in accounting (Khumawala & Gordon, 1997; Tinkelman, 1998, 1999), four studies in economics (Okten & Weisbrod, 2000; Steinberg, 1986a, b; Weisbrod & Dominguez, 1986) and one study in public administration (Frumkin & Kim, 2001) examine effects of information derived from NPO financial disclosures on private donations to U.S. NPOs. All of these studies, except Khumawala and Gordon (1997) and Frumkin and Kim (2001), use models that include one factor, termed "price" of donating, calculated from required NPO financial disclosures.[1] The economics studies of U.S. NPOs specify "price" as 1/(1 – "fundraising expenses"/"total private donations").[2] The accounting studies specify "price" as the reciprocal of "administrative efficiency." "Administrative efficiency" is "program expenses"/"total expenses," which are line items on NPO financial statements.

"Watchdog" agencies (e.g. National Charities Information Bureau and American Institute of Philanthropy) and the popular press (e.g. *Money* and *Fortune*) publish "administrative efficiency." No study tests a model that includes NPO wealth.

Khumawala and Gordon (1997) use an experimental design to test the effect of administrative efficiency provided to subjects at no cost. This design provides evidence that donors, provided with the administrative efficiency of NPOs, incorporate administrative efficiency in their donation decisions. However, the design does not provide evidence on whether donors, subject to search costs in their real-world donation decisions, incorporate administrative efficiency. Tinkelman (1998, 1999) tests cross-sectional data on 151–178 large U.S. NPOs, and 6,500 U.S. NPOs, respectively, in levels form. These studies implicitly test whether donors actually incorporate administrative efficiency in their real-world donation decisions. However, models in levels form do not control for unspecified omitted organization-specific factors that are correlated with one or more independent variables in the model.Weisbrod and Dominguez (1986) apply neoclassical economic theory of consumer demand in modeling aggregated private donations to a particular NPO as a function of the organization-specific factors price, advertising, and quality. They use 1/(1 – "fundraising expenses"/"total private donations") to proxy for price, fundraising expense to proxy for advertising, and age of the organization to proxy for quality. Weisbrod and Dominguez lag fundraising one year to reflect their view that "potential donors subjected to fundraising in one period respond in the following period if at all." Weisbrod and Dominguez lag price one year, assuming price lagged one year is available to donors for the entire fiscal year in which they donate. Because of data limitations, in specifying price, Weisbrod and Dominguez assume that "administration expenses" are zero. Because they examine the elasticity of donations to fundraising expenses and implicitly assume that donors think their donations are used differently than other revenues, they use "private donations" instead of "total expenses." This results in a specification of price that is nearly orthogonal (Pearson correlation of 0.026 and Spearman correlation of 0.164) to the price used in the current study and the other accounting studies. Steinberg (1986a, b) specifies price in the same way as Weisbrod and Dominguez.

Posnett and Sandler (1989) expand Weisbrod and Dominguez's model to include government support and "autonomous income" (NPO rental and investment income, and fees charged for services), to test "crowding-out" of private donations by these other types of revenue. They specify price as the reciprocal of "administrative efficiency." Posnett and Sandler test 300 large U.K. NPOs and find a significantly negative relation between price and private donations.

Tinkelman (1998) adds total assets of the NPO to the Posnett and Sandler model. He tests data on large U.S. NPOs in levels form and finds a significant negative relation between price (the reciprocal of "administrative efficiency") and

private donations. Tinkelman (1999) tests the same model using New York State regulatory filings data on 6,500 U.S. NPOs for 1993 and 1994 divided into nine industry sub-samples, in levels form. He finds that price is negatively related to donations in all nine industry samples. However, he also states that results from testing the model in first-differences form indicate no significant relation between price and donations, and the coefficient of determination of this model is only 0.01.

Okten and Weisbrod (2000) test a model similar to Posnett and Sandler's in levels form using data for U.S. NPOs in seven industries and then using two stage least squares regression in first differences form to test panel data in three industries – hospitals, higher education, and scientific research. They specify price in the same way as Weisbrod and Dominguez (1986).

Frumkin and Kim (2001) test a one-way random effects model using the same panel data set used in the current study. They find no significant relation between price and private donations for any of the six industry samples they test, and from this conclude that NPO efficiency has no effect on private donations. However, their specification of price is quite different from the specification used in the prior studies; they ignore fundraising expense. Furthermore, their model departs significantly from models used in prior studies in that it does not include age, program service revenue (or wealth) and introduces two new variables, total revenue and program expense, without explanation. Therefore, their results are not comparable to the results of the current study or other prior studies.

3. THEORY

The American Institute of Philanthropy's (AIP) *Charity Rating Guide* reports the statistic "years of available assets." This statistic, calculated from line items on the NPO financial statements, is

$$Y = \frac{\text{NA}}{\text{TEXP} - \text{FR}} \tag{1}$$

where NA is total net assets less permanently restricted net assets at the beginning of the current year, TEXP is total expenses of the preceding year, and FR is fundraising expenses of the preceding year. AIP describes this statistic as "how long a charity could continue to operate at current levels without additional fundraising" and characterizes this as a measure of NPO "wealth."

Anecdotal evidence suggests NPO management believes the appearance of "wealth" is undesirable. For example, the NPO Boys Town sued a "watchdog" agency for characterizing it as one of the wealthiest charities (based on its years of available assets) claiming that the characterization reduced donations by two

million dollars (Sebastian, 1994). According to interdependent utility theory developed by Schwartz (1970) and Becker (1974), an individual can increase her utility by increasing certain other persons' consumption of a particular good. A primary way to do this is by donating to an NPO that produces the good, thereby increasing the output of the NPO. Donors should obtain greater utility from others' current consumption of goods provided by the NPOs to which they donate than from others' future consumption of such goods. Thus, a dollar donated to an NPO that takes less time to convert the dollar to output provides more utility to a donor, ceteris paribus, than a dollar donated to an NPO that takes more time to convert the dollar to output.[3] In other words, an individual's utility from his donation should decrease as the amount of time the individual perceives it takes the NPO to produce output from the current donation increases, ceteris paribus. Information contained in the financial statements may provide information to donors on the amount of time it takes an NPO to produce output from a current donation.

The number of years that donors might perceive, during year *t*, that it will take the NPO to spend its current surplus and, thus, be ready to produce output from donations received in year *t*, is

$$T_t = \frac{\mathrm{NA}_t}{\mathrm{TEXP}_{t-1}} \tag{2}$$

where NA_t is net assets (total net assets less permanently restricted net assets) at the beginning of year *t*, and TEXP_{t-1} is total expenses of year $t-1$. A negative relation between an individual's utility from donating, and therefore, amount of donations, and *T* is hypothesized.

The difference between *T* and *Y* (in Eq. (1)) is that fundraising expense is subtracted in the denominator in calculating *Y*. While *T* and *Y* clearly are not equal, empirically *T* and *Y* are highly correlated (0.98) in the ten years of data used in this study. Therefore, results from testing *Y* and *T* in a model would be very similar. In other words, while theory does not directly support introducing *Y* into a model of private donations, empirically there should be little difference between testing *T* and *Y*, and *Y* is the statistic published by the "watchdog" agency, AIP.[4]

4. EMPIRICAL SPECIFICATIONS

Based on the previous discussion, "years of available assets" is added to Tinkelman's (1999) model to produce the following model.

$$\begin{aligned}\ln\mathrm{DON}_{i,t} = {} & b_0 + b_1 \ln\mathrm{PRICE}_{i,t-1} + b_2 \ln Y_{i,t} + b_3 \ln\mathrm{FR}_{i,t-1} + b_4 \ln\mathrm{GOV}_{i,t-1} \\ & + b_5 \ln\mathrm{PREV}_{i,t-1} + b_6 \ln\mathrm{AGE}_{i,t} + b_7 \ln\mathrm{TOTASS}_{i,t} + u_{i,t}\end{aligned} \tag{3}$$

where i indicates NPO, t indicates year, DON is private (non-governmental) donations, PRICE is total expenses divided by program expenses; i.e. reciprocal of administrative efficiency, Y is years of available assets[5] at the beginning of the year, FR is fundraising expense, GOV is government contributions, PREV is program service revenue,[6] AGE is organization age; i.e. time since founding, in years, TOTASS is total assets at the beginning of the year, and u is error.[7]

The model is tested in levels form using OLS regression on pooled data to compare results to those of prior research. The model is then reparameterized to a one-way error components model using panel data. One-way error components models control for unspecified organization-specific factors, such as inherent attractiveness of NPO mission, size of NPO constituency, etc. that are essentially constant over time and that may be correlated with the independent variables. Estimates of parameters from panel data are more persuasive than estimates from time-series or cross-sectional data (Barrett, 1991).[8]

To incorporate unspecified NPO-specific effects (that are inter-temporally constant), the error term of the cross-sectional model is rewritten as:

$$u_{i,t} = m_i + e_{i,t} \tag{4}$$

where m_i is the effect of unspecified cross-sectionally variant but inter-temporally constant factors, and $e_{i,t}$ is the remaining error.

The term m_i depends only on the class, i, to which the observation belongs, and the effect of this variable is incorporated into the organization-specific intercept term. Thus, Eq. (3) can be rewritten as:

$$\begin{aligned}\ln\mathrm{DON}_{i,t} = {} & b_{0i} + b_1 \ln\mathrm{PRICE}_{i,t-1} + b_2 \ln Y_{i,t} + b_3 \ln\mathrm{FR}_{i,t-1} + b_4 \ln\mathrm{GOV}_{i,t-1} \\ & + b_5 \ln\mathrm{PREV}_{i,t-1} + b_6 \ln\mathrm{AGE}_{i,t} + b_7 \ln\mathrm{TOTASS}_{i,t} + u_{i,t}\end{aligned} \tag{5}$$

where b_{0i} is the intercept particular to the ith NPO.

Error components models can be classified into two types – fixed effects and random effects, depending on whether the sample is randomly selected from a population (Hsiao, 1986). Since the sample is not randomly selected, a fixed effects model is used. Thus, the results are applicable only to the organizations in the sample; i.e. large NPOs (those with total assets of $10 million or more), and are not generalized to the population of all NPOs. To determine whether the pooled cross-sectional or one-way fixed effects is the "better" model, a simple F test is used (Barrett, 1991). The F test indicates that the one-way fixed effects model is the "better" model in every industry sample.

5. DATA

The National Center for Charitable Statistics Statement of Income database provides data for all variables in this study, except age. The underlying source of this database is Internal Revenue Service (IRS) Forms 990. The database contains data for all U.S. NPOs for years 1985–1994 with assets at least $10 million in a given year and for a random sample of NPOs with assets less than $10 million in a given year. Since developing a set of panel data requires data for the same organizations across all years, the random sample of NPOs with assets less than $10 million cannot be included in a set of panel data. Thus, the data tested are for large NPOs, those with assets of $10 million or more, in every year from 1985 to 1994. The initial panel data set consists of 1,950 large NPOs for a ten-year period.

Tinkelman (1998, 1999) specify age as the number of years the organization has registered with New York state, a practice that began in 1954. He does this because these data are readily available from the state filing database he uses. Specifying age in this way necessarily truncates age at 40, and biases parameter estimates for age downward and significance tests against finding significance. The current study specifies age as actual age of the NPO since founding, which is an untruncated measure. These data were collected from various sources. *American Universities and Colleges* (Murray, 1999) provides age for universities and colleges, *The Handbook of Private Schools* (1999) provides age for other schools, *The Official Museum Directory* (2000) provides age for museums, and the *Encyclopedia of Associations* (Craddock & Fisher, 1996) provides age for some other NPOs. For many NPOs, age is obtained from their website or directly from a knowledgeable employee of the NPO. However, data on age cannot be obtained for 675 NPOs because the NPO is not found in any of the references nor can it be contacted. Thus, from a beginning set of panel data of 1,950 NPOs, 675 NPOs are eliminated because data on age cannot be obtained. Furthermore, an additional 36 NPOs are deleted from the panel data set because one or more of the variables for the NPO is implausibly negative.[9] This leaves a panel data set of 1,239 NPOs for the ten-year period 1985–1994. However, since some variables are lagged one year, only nine years of data, or 11,151 observations in total, are usable.

Since the log of zero is undefined, following the prior research, a nominal amount ($1) is added to every zero value in the data. To reduce the impact of extremely large values for price, which can be infinite as defined, price is winsorized at 10 following Tinkelman (1999). Similarly, to reduce the impact of extremely large values of wealth, wealth is winsorized at 50.[10] Furthermore, to reduce measurement error from nominal changes in monetary variables due to inflation, all monetary variables are deflated by the Consumer Price Index.

Prior research (e.g. Okten & Weisbrod, 2000; Weisbrod & Dominguez, 1986) finds significant differences in results for NPOs in different industries (e.g.

education, health, and arts).[11] Therefore, the following mutually exclusive and collectively exhaustive industry samples are tested: arts, education, health, human services, philanthropic, and "other" organizations. Data limitations preclude breaking the "other" industry sample into further industry samples. NPOs are classified into one of these industry samples based on the National Tax Exempt Entity classification scheme (provided in the data). Descriptive statistics for each industry sample are provided in Table 1.

Table 1. Descriptive Statistics.

	Arts	Education	Health	Human Services	Philanthropic	Other
Donations $ in (000s)	11,757	9,744	2,429	6,070	21,847	14,521
	6,114	3,653	300	1,288	12,246	3,304
	23,125	21,424	9,863	17,094	29,602	38,347
Price	1.68	1.27	1.35	1.73	1.14	1.43
	1.28	1.17	1.16	1.15	1.12	1.22
	4.13	0.53	1.04	2.09	0.12	1.11
Years of available assets	5.06	3.07	1.70	4.92	6.03	3.91
	4.02	2.16	0.54	1.66	3.79	1.58
	3.88	3.16	6.00	9.68	6.34	6.45
Fundraising $	881	1,019	139	248	1,172	1,535
	517	525	0	0	247	190
	1,485	1,630	813	677	2,442	4,103
Government support $	2,876	9,825	1,479	4,877	468	8,774
	532	880	0	0	0	0
	6,757	42,358	9,956	30,891	2,261	45,200
Program revenue $	9,238	42,256	109,782	9,303	420	16,708
	2,784	14,260	70,786	5,371	0	1,599
	22,722	91,652	268,517	12,320	1,892	43,882
Age (years)	74	109	77	86	65	64
	74	106	79	90	67	55
	35	49	38	46	24	39
Total assets $	98,200	178,549	131,711	84,147	82,713	76,997
	65,280	65,893	81,694	35,929	60,960	50,694
	106,911	441,443	216,192	186,290	87,239	80,784
Sample size	549	4,086	4,662	495	441	918

Note: Descriptive statistics (mean, median, and standard deviation) for a panel data set of large U.S. nonprofit organizations from 1986 to 1994. Statistics are for each of six industry samples, based on the National Tax Exempt Entities classification scheme. Donations, fundraising, government support, program revenue, and total assets are in thousands of dollars. Price is total expenses/program expenses, which is the reciprocal of "administrative efficiency."

Because of significant heteroscedasticity in all years, White's (1980) consistent variance-covariance matrix estimator is used to develop confidence intervals. Multi-collinearity, measured by condition indices is moderate. Cook's distance test indicates no influential outliers in any of the industry samples. Low Durbin-Watson *d* statistics indicate significant positive autocorrelation in all samples. To correct for autocorrelation, the Cochrane – Orcutt iterative method is used (Greene, 2002).

6. RESULTS FROM TESTING POOLED CROSS-SECTIONAL DATA

Results from OLS regression using pooled cross-sectional data for each industry sample are presented in Table 2.

The fully robust significant results of OLS regression using pooled cross-sectional data of this study, and results of Tinkelman (1998, 1999) (the only comparable studies of U.S. NPOs) appear in Table 3. Tinkelman (1998) reports results only for a full sample in each year; he tests no industry samples. Tinkelman (1999) reports results only for a full sample for each year for all variables except price. He reports the price variable by industry sample. However, he uses data from New York State regulatory filings and, therefore, bases his industry samples on the New York state classification system. He also removes hospitals and schools from his health and education industry samples. The industry samples in the current study are based on the NTEE classification codes, and hospitals and schools are not deleted. Therefore, the results are not directly comparable. Nonetheless, the only significant fully robust estimate (for education NPOs) of the price coefficient in the current study is within the range of estimates for price from Tinkelman (1999). The estimate for the price coefficient for the arts sample (shown in Table 2) is positive, but not significant. However, when years of available assets, the independent variable newly introduced into the model in this paper, is omitted from the model, the estimated price coefficient for the arts sample becomes significantly negative (−0.41 at the 1% level), consistent with Tinkelman's results. Estimated coefficients for all variables for all industry samples, except the price coefficient for the arts sample, are robust to omitting years of available assets. However, the change in estimated price coefficient for the arts sample, from insignificantly positive to significantly negative, suggests that years of available assets may be an omitted correlated variable in Tinkelman's model and the other models used in prior studies.

Estimates for fundraising are generally smaller in the current study. Estimates for age are generally substantially higher in the current study than in Tinkelman

Table 2. Pooled Cross-sectional Results.

	Arts ($n = 549$)	Education ($n = 4{,}086$)	Health ($n = 4{,}662$)	Human Services ($n = 495$)	Philanthropic ($n = 441$)	Other ($n = 918$)
INTERCEPT	4.20*	0.08	−2.00	21.7*	−10.2*	7.14**
t stat.	5.9	0.1	−1.4	8.2	−2.8	2.4
PRICE	0.13	−0.70*	0.07[b]	−0.74[a,**]	2.72[b]	−0.06[a,b,c]
t stat.	0.96	−3.7	0.3	−2.6	1.5	−0.1
Y	−0.50*	0.48*	0.43[a,*]	0.11[b]	−0.75*	0.26[a,b,**]
t stat.	−10.6	11.2	5.0	1.2	−6.8	2.3
FR	0.03*	0.18*	0.15*	0.19*	0.12*	0.38*
t stat.	3.9	20.8	12.4	7.7	4.2	18.0
GOV	0.05*	0.07*	0.12*	–0.06**	0.03[b]	0.06[a,*]
t stat.	4.9	10.1	11.3	–2.5	1.2	3.0
PREV	−0.05*	−0.08*	−0.01	−0.28*	−0.01[b]	−0.00[b]
t stat.	−5.7	−7.5	−0.3	−11.0	−0.5	−0.2
AGE	−0.21[a,*]	1.83*	0.72*	0.76*	1.92[b,*]	1.20*
t stat.	−3.4	24.3	6.9	4.3	5.6	4.5
TOTASS	0.70*	0.24*	0.54*	−0.48[b,*]	0.96*	−0.11[b]
t stat.	17.3	7.1	6.7	−3.6	4.9	−0.7
Adj. R^2	0.53	0.46	0.13	0.35	0.33	0.37

Note: Regression coefficients from testing the following model in levels form using pooled cross-sectional data for each of six industry samples based on the National Tax Exempt Entities classification scheme.

$$\ln\text{DON}_{i,t} = b_0 + b_1 \ln\text{PRICE}_{i,t-1} + b_2 \ln Y_{i,t} + b_3 \ln\text{FR}_{i,t-1} + b_4 \ln\text{GOV}_{i,t-1} + b_5 \ln\text{PREV}_{i,t-1} + b_6 \ln\text{AGE}_{i,t} + b_7 \ln\text{TOTASS}_{i,t} + u_{i,t}$$

DON is private donations (in dollars). PRICE is price (the reciprocal of "administrative efficiency") which is "total expenses"/"program expenses." *Y* is "years of available assets" specified as "net assets"/("total annual expenses" – "fundraising expenses"). FR is fundraising expense (in dollars). GOV is governmental financial support (in dollars). PREV is program revenue (in dollars). AGE is age of the organization (in years). TOTASS is total assets at the beginning of the year (in dollars).

[a] Not robust to specifying *Y* as ("net assets" – "fixed assets")/("total expenses" – "fundraising expenses").

[b] Not robust to rank regression.

[c] Not robust to winsorizing price at 22,000.

* Significant at the 1%, levels (two-tailed).

** Significant at the 5%, levels (two-tailed).

Table 3. Comparative Results.

Study	Sample	PRICE	FR	AGE	PREV	GOV	TOTASS
This study	1,239 large U.S. NPOs for 1986–1994	−0.70	0.03 to 0.38	0.72 to 1.83	−0.05 to −0.28	−0.06 to 0.12	0.24 to 0.96
Tinkelman (1998)	151–178 U.S. NPOs for 1991–1992	−0.89 and −1.48	0.58 and 0.59	−0.14 and −0.14	−0.02 and −0.03	-	0.27 and 0.29
Tinkelman (1999)	6,500 U.S. NPOs for 1993–1994	−0.19 to −1.17	0.52 and 0.53	0.07 and 0.09	−0.08 and −0.08	−0.19 and −0.20	0.26 and 0.27

Note: A comparison of fully robust regression results of this study and prior studies using cross-sectional or pooled cross-sectional data.

The dependent variable in all studies is aggregated private donations to a given NPO. Numerical entries are elasticities and are significant at the 10% level or better. GOV is not significant in Tinkelman's (1998) study.

Cells for all variables for this study contain the range of fully robust significant coefficient estimates across all industry samples. The cell for PRICE for Tinkelman's (1999) study contains the range of significant coefficient estimates across all industry samples. Cells for all variables for Tinkelman's (1998) study and all variables, except PRICE, for Tinkelman's (1999) study contain coefficient estimates for the full samples for each of the two years of the respective studies.

PRICE is "total expenses"/"program expenses" (reciprocal of "administrative efficiency"), FR is fundraising expense (in dollars), AGE is age of the organization (in years), PREV is program service revenue (in dollars), GOV is governmental financial support (in dollars), TOTASS is total assets at the beginning of the year (in dollars).

(1998, 1999). This could be because the current study specifies age with less measurement error than the Tinkelman studies. The current study specifies age as years since the NPO was founded, whereas the Tinkelman studies specify age as years since the NPO registered with New York state, which truncates the data.

7. RESULTS FROM TESTING THE ONE-WAY FIXED EFFECTS MODEL USING PANEL DATA

Results shown in Table 4 are for the one-way fixed effect model using industry samples.

The results are substantially different from the results of the levels models using pooled cross-sectional data (shown in Table 2). However, while using error components models has the advantage of reducing inconsistency from correlated omitted variables, it has the disadvantage (along with differences models) of exacerbating the effects of measurement error, the so-called "errors in variables" problem (Hsiao, 1986). In addition, the effect of age in one-way error components models is independent of the actual age of the NPO. Whether the NPO is two years old or 200 years old, the effect of age is implicitly assumed to be the same because one-way error components models use the difference from the mean.

Griliches and Hausman (1986) suggest that one can obtain evidence of the seriousness of measurement error in variables by testing the data using various differences forms of the model and comparing parameter estimates across the various differences models. Wide differences in estimates for a given variable suggest serious measurement error in the variable. Each industry sample is tested using first-differences models for each pair of years in the panel data, and the parameter estimates across the differences models are compared. Wide differences in the parameter estimates, including different signs, are found in all industry samples for the variables fundraising, government support, program service revenue, and age. Wide differences, including change in sign, are also found for the price variable in all industry samples except arts. This suggests substantial measurement error in these variables (Griliches & Hausman, 1986), implying that parameter estimates are biased toward zero and significance tests are biased against significance.

Unfortunately, substantial measurement error in some of the variables contaminates results, even if significant, for the other variables (Greene, 1993). Thus, no results from error components are discussed. The important conclusion from these tests is that the NCCS panel database appears to contain too much measurement error to test effectively error components models.

Table 4. One-way Fixed Effects Model Results – Industry Samples.

	Arts ($n = 549$)	Education ($n = 4086$)	Health ($n = 4662$)	Human Services ($n = 495$)	Philanthropic ($n = 441$)	Other ($n = 918$)
PRICE	0.41***	0.05*	0.03	−0.70***	−0.38	0.10
t stat.	1.7	0.5	0.1	−1.7	−1.6	0.5
Y	−0.53*	−0.11*	−0.41***	0.02	−0.27*	−0.24
t stat.	−5.0	−2.9	−1.7	−0.1	−2.9	−1.2
FR	0.01	0.00	0.02	0.01	0.01	0.01
t stat.	0.8	0.5	1.2	0.2	0.6	0.2
GOV	0.00	−0.00	−0.00	−0.04**	−0.00	0.01
t stat.	0.1	−0.4	−0.1	−2.1	−0.2	0.3
PREV	−0.02	0.00	−0.02	−0.29**	0.02	−0.01
t stat.	−1.3	0.1	−0.7	−2.0	0.2	−0.8
AGE	−0.04	0.55	0.45	1.92	0.30	1.95
t stat.	−0.1	1.3	0.5	1.1	0.7	1.0
TOTASS	0.56*	0.09	0.09	0.26	0.65*	1.17**
t stat.	4.0	1.0	0.3	0.4	3.1	2.0
Adj. R^2	0.88	0.97	0.67	0.73	0.99	0.89

Note: Regression coefficients from testing the following model, specified as one-way fixed effects using panel data for each of six industry samples based on the National Tax Exempt Entities classification scheme

$$\ln\text{DON}_{i,t} = b_{0i} + b_1 \ln\text{PRICE}_{i,t-1} + b_2 \ln Y_{i,t} + b_3 \ln\text{FR}_{i,t-1} + b_4 \ln\text{GOV}_{i,t-1} + b_5 \ln\text{PREV}_{i,t-1} + b_6 \ln\text{AGE}_{i,t} + b_7 \ln\text{TOTASS}_{i,t} + u_{i,t}$$

DON is private donations (in dollars). PRICE is price (the reciprocal of "administrative efficiency") which is "total expenses"/"program expenses." *Y* is "years of available assets" defined as "net assets"/("total annual expenses" – "annual fundraising expenses"). FR is fundraising expense (in dollars). GOV is governmental financial support (in dollars). PREV is program service revenue (in dollars). AGE is age of the organization (in years). TOTASS is total assets at the beginning of the year (in dollars).

* Significant at the 1%, levels (two-tailed).

** Significant at the 5%, levels (two-tailed).

*** Significant at the 10%, levels (two-tailed).

8. EFFECT OF WEALTH (YEARS OF AVAILABLE ASSETS)

An NPO's wealth, proxied by the financial disclosure "years of available assets," is hypothesized to have a negative effect on donations. Results for two of the industry

samples, arts and philanthropic NPOs (shown in Table 2) support this hypothesis; years of available assets (Y) is negatively related to private donations at the 1% level. However, results for education NPOs are inconsistent with the hypothesis; years of available assets is *positively* related to private donations at the 1% level. Results for the other industry samples, health, human services, and "other," are not discussed because they are not fully robust (see Notes 4 and 10).

One explanation for a *positive* relation between private donations and years of available assets is that higher years of available assets allows NPOs to spend more on fundraising, thereby increasing private donations. However, the correlation of years of available assets with fundraising for education NPOs is low (−0.03). Thus, this explanation is unlikely to be valid.

Another explanation for the significant positive relation between years of available assets and private donations is more complex and relates to NPO accounting conventions. NPOs may receive private donations in the form of endowments. NPOs record endowments as private donation revenue in the year received, and typically invest endowments because they may not spend endowments, but may spend the returns from investing endowments. Thus, an endowment increases total donations of an NPO, and, ceteris paribus, increases assets without increasing liabilities, thereby increasing net assets (more specifically, permanently restricted net assets). It is important to note that the theory developed in this paper for the relation between wealth and private donations is not applicable to endowments; donors who provide endowments know that the NPOs will never spend the endowment. This suggests that donors of endowments may receive a different kind of utility from providing endowments that more than offsets the theorized disutility stemming from the NPOs' permanently postponed spending of the endowment. Thus, NPO wealth might not negatively affect donors who provide endowment donations.

Because an NPO may spend the returns on investing the endowment, but not the endowment itself, in the year the NPO receives an endowment, it may increase its expenses up to the amount of the return on the invested endowment, without additional borrowing. Therefore, in the year an NPO receives an endowment, the NPO's net assets necessarily increase and its expenses may increase. Thus, an NPO's years of available assets as specified in this study ("beginning of year net assets"/("prior year total expenses" – "prior year fundraising expenses")) might also increase, ceteris paribus, in the year the NPO receives an endowment simply because the NPO received an endowment donation. This suggests a positive relation between endowment donations and years of available assets (a proxy for wealth) as specified in this study.

It is conceivable that this positive relation between endowment donations and years of available assets for education NPOs overwhelms the negative relation

between non-endowment private donations to large education NPOs and years of available assets. It is also conceivable that this positive relation between endowment donations and years of available assets for arts and philanthropic NPOs does not overwhelm the negative relation between non-endowment private donations to arts and philanthropic NPOs and years of available assets. This would explain the empirical result of a positive relation between total private donations (endowment and non-endowment donations) and years of available assets for large education NPOs, and a negative relation between total private donations (endowment and non-endowment) and years of available assets for large arts and philanthropic NPOs. To test the validity of this explanation, one could exclude endowment donations from the specification of total private donations and exclude permanently restricted net assets from the specification of years of available assets. Unfortunately, data on endowment donations and permanently restricted net assets are not included in the NCCS panel data set used in this study, and these data are not readily available elsewhere. Therefore, providing evidence on the validity of the preceding explanation is left to a future study.

Another explanation for the positive relation between wealth and donations to education NPOs is as follows.[12] Direct beneficiaries of the output of education NPOs in year *t*, namely students, are also a primary source of future donations as alumni, say in year $t + 10$ or $t + 20$. Moreover, it is reasonable that the greater the quality and quantity of output of a given education NPO in year *t*, the greater the capacity and willingness of alumni who consumed output in year *t* to donate in the future, say at time $t + 10$ or $t + 20$. If wealth of an education NPO is positively related to the quality and quantity of its output, and wealth of an education NPO in year *t* is positively correlated with its wealth in year $t + 10$ or $t + 20$, then alumni donations to the NPO in time $t + 10$ or $t + 20$ should be positively related to its wealth. This phenomenon is likely to be limited to education NPOs. This is because the quality and quantity of an education NPO's output should have a greater effect on the donative willingness and capacity of its beneficiaries, namely students, than the quality and quantity of other types of NPOs' output has on the donative willingness and capacity of their beneficiaries.

9. CONCLUSIONS

I find evidence of substantial measurement error in the variables of fundraising, government support, program service revenue, and age, in the one-way fixed effects model. Unfortunately, serious measurement error in some of the variables contaminates results, even if significant, for the other variables (Greene, 1993). Thus, the conclusion drawn from testing the one-way fixed effects model is that

the NCCS panel database, a recently developed database that is "expected to be the tax-exempt sector's equivalent to the Compustat databases . . ." (Gordon et al., 1999, p. 113), may be limited in its usefulness due to measurement error.

Results for the pooled cross-sectional model, where measurement error is not as serious as in the one-way fixed effects model, show that age has a major impact on private donations to large NPOs (those with total assets of $10 million or more). This study obtained hand-collected data on age of large NPOs since founding. This is a better specification of age than "years since registering with the state," a proxy used in prior studies that truncates the age variable. The significant and fully robust age coefficients range from 0.72 to 1.83 for the education, health, human services, and "other" industry samples. These results suggest that older NPOs in these industries receive, on average, more private donations than younger NPOs, presumably because over time they develop a greater stock of goodwill with the public.

I also provide evidence on the effects of a proxy for the wealth of an NPO, "years of available assets," on private donations to large U.S. NPOs. "Years of available assets" is calculated from NPO financial disclosures. As theory predicts, NPO wealth is found to have a significant negative effect on private donations to large arts and philanthropic NPOs. These results support managers and directors of large NPOs in these industries who believe that an NPO's wealth negatively impacts private donations. However, I also find a positive relation between wealth and private donations to large education NPOs. A proposed explanation for this result is that the positive relation between endowment donations and wealth of education NPOs, stemming from the way NPOs account for endowment donations, might more than offset the hypothesized negative relation between wealth and non-endowment private donations to education NPOs.

The coefficient estimate for "price" (the reciprocal of "administrative efficiency," a proxy for NPO efficiency) in the pooled cross-sectional regression model is positive, but not significant, for large arts NPOs. However, when years of available assets, the independent variable newly introduced into the model in this paper, is omitted from the model, the price coefficient for the arts sample becomes significantly negative (-0.41 at the 1% level), consistent with Tinkelman's (1999) results. This suggests that years of available assets may be an omitted correlated variable in Tinkelman's model and the other models used in prior studies. The estimates of the price coefficient for all other types of large NPOs, except for education NPOs, is not robust to three robustness tests. Private donations to large education NPOs are negatively (positively) related to the price (administrative efficiency) of such NPOs. The coefficient estimate for price is -0.70, consistent with results of prior studies, and an indication that NPO efficiency has a large significant positive effect on private donations to large education NPOs.

NOTES

1. Khumawala and Gordon (1997) use an experimental method to test two factors from NPO financial disclosures. One factor is "administrative efficiency," similar to the other accounting studies. The other factor is non-monetary information provided in a statement of service efforts and accomplishments, which is not a required disclosure for NPOs. Frumkin and Kim (2001) use a factor they term "efficiency," which is "administration expenses"/"total expenses" from NPO financial disclosures.

2. The economics studies develop theory that motivates including the reciprocal of administrative efficiency, which they term "price," in these models. One of the components of price is "total expenses" comprised of the mutually exclusive and collectively exhaustive components "fundraising," "administration," and "program." However, because of data limitations, the economics studies of U.S. NPOs assume that "administration" is zero. Furthermore, they use "total private donations" instead of "total expenses" in their specification of price under the implicit assumption that donors think that donations are used differently than other revenues.

3. An NPO that does not immediately spend a donated dollar presumably would invest it. However, it is plausible that donors' discount rates implicit in the difference in utility obtained from the NPO spending the donated dollar currently, instead of in the future, is greater than donors' expectation of the NPO's return on investing the donated dollar. On the other hand, some donors provide donations in the form of endowments, which NPOs may never spend. NPOs may spend only the returns from investing endowments. The theory developed in this paper does not explain endowment donations, which might not be affected by wealth in the same way as non-endowment donations.

4. It is plausible that donors with a sophisticated understanding of accounting would refine the AIP's specification of "wealth" by subtracting fixed assets from net assets. This refined specification would distinguish NPOs with high wealth in the form of fixed assets from NPOs with high wealth from unrestricted cash and other current assets. I test the robustness of results to this revised specification; i.e. ("net assets" – "fixed assets")/("total expenses" – "fundraising expenses"), in the pooled cross-sectional models. All results are robust to this specification, except for the following: Age in the arts sample is no longer significantly negative. Years of available assets in the health sample is no longer significant. Price in the human services sample is no longer significantly negative. Price in the "other" sample goes from significantly negative to significantly positive, and years of available assets and government funding in the "other" sample are no longer significantly positive. These robustness results are indicated in Table 2. Results that are not robust are considered inconclusive and are not interpreted.

5. The source for the data set used in this study, the IRS Form 990, does not provide information on "permanently restricted net assets." This introduces measurement error and may drive the result of a positive relation between wealth and private donations. (See Section 9 of the paper for a detailed explanation.) Thus, in this study, "net assets" is specified as "total net assets," without subtracting "permanently restricted net assets." However, the IRS Form 990 is a major source of information for potential donors, and it is the only information that NPOs are required, by law, to make available to the public. Furthermore, donors with a sophisticated understanding of accounting might refine the calculation of years of available assets by subtracting fixed assets from net assets. See note 4 for further discussion

and results of testing the robustness of results to this refined specification of years of available assets.

6. PREV, a component of Posnett and Sandler's "autonomous income" that is never negative, replaces "autonomous income" because some values of the latter are negative and the log of a negative number is undefined. One may add a constant to negative values to make them positive, but only when the hypothesized relation is monotonic from negative to positive values of the variable. The economics literature hypothesizes the relation of autonomous income to donations not to be monotonic from negative values to positive values of autonomous income.

7. The interaction term AGE × FR, included in some models, was also included in the original model and tested in levels form. However, it introduced excessive multicollinearity, and, therefore, was dropped. Tinkelman's (1999) model does not include this term, and he does not discuss whether it was tested.

8. However, error components models can exacerbate measurement error in variables, and thus bias estimates of parameters toward zero and significance tests against finding significance. Results from applying the method of Griliches and Hausman (1986) indicate substantial measurement error in the variables fundraising, government support, program revenue, and age in all industry samples. Attempts to quantify the measurement error and obtain consistent parameter estimates, using the method of Hsiao (1986), generally produced meaningless results.

9. Wealth plausibly could be negative because net assets (or alternatively net assets – fixed assets, see Note 4) plausibly could be negative; i.e. the NPO has greater liabilities than assets. Theory suggests that donations are negatively related to wealth. However, increasingly negative wealth should not increase donations. The relation of wealth to donations may be quadratic. The model using a squared term for wealth was tested, but this term introduced excessive multicollinearity, and therefore, was dropped. Thus, negative values of wealth must be eliminated from the sample, and since it is beneficial to retain a balanced panel, NPOs with negative wealth in any year are deleted.

10. As a robustness test, price is winsorized at 22,000 and wealth is winsorized at 1,096. Results are fully robust to this test, except for price in the "other" industry sample, which becomes significantly negative. As an additional robustness test, rank regression is performed. Results for the arts and education samples are fully robust. However, for the health sample, price goes from insignificantly positive to significantly positive. For the human services sample, years of available assets goes from insignificantly positive and very small to significantly negative, and total assets goes from significantly negative to significantly positive. For the philanthropic sample, price goes from insignificantly positive, but large, to nearly significantly negative; government spending goes from insignificant to significantly positive, program service revenue goes from insignificant to significantly negative, and age is no longer significantly positive. For the "other" sample, price is no longer significantly negative, years of available assets goes from significantly positive to significantly negative, program service revenue becomes significantly negative, and total assets becomes significantly positive. All of the robustness results are indicated in Table 2. All results that are not robust to rank regression are considered inconclusive and are not interpreted.

11. Applying a Chow test to the data in the current study to test the null hypothesis of equality of the parameters across all industry samples rejects the null at the 1% level; i.e. parameters are not equal across industry samples.

12. The author thanks the anonymous reviewer for suggesting this explanation.

ACKNOWLEDGMENTS

I have benefited greatly from suggestions from Fred A. Jacobs, Dave Laband, John Sophocleus, and Ping Zhou. I am very appreciative of the many substantive suggestions made by the anonymous reviewer.

REFERENCES

AAFRC Trust for Philanthropy (2001). *Giving USA*. IN: AAFRC Trust for Philanthropy.

Barrett, K. S. (1991). Panel-data estimates of charitable giving: A synthesis of techniques. *National Tax Journal*, *44*(September), 365–381.

Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy*, *82*(61), 1063–1093.

Craddock, J., & Fisher, C. (Eds) (1996). *Encyclopedia of associations – National organizations of the U.S*. New York, NY: Gale Research.

Frumkin, P., & Kim, M. T. (2001). Strategic positioning and the financing of nonprofit organizations: Is efficiency rewarded in the contributions marketplace? *Public Administration Review*, *61*(3), 266–275.

Gordon, T. P., Greenlee, J. S., & Nitterhouse, D. (1999). Tax-exempt organization financial data: Availability and limitations. *Accounting Horizons*, *13*(2), 113–128.

Greene, W. (1993). *Econometric analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Greene, W. (2002). *LIMDEP econometric modeling guide*. Plainview, NY: Econometric Software.

Griliches, Z., & Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, *31*, 93–118.

Hsiao, C. (1986). *Analysis of panel data*. New York, NY: Cambridge University Press.

Khumawala, S. B., & Gordon, T. P. (1997). Bridging the credibility of GAAP: Individual donors and the new accounting standards for nonprofit organizations. *Accounting Horizons*, *11*(3), 45–68.

Murray, J. J., III (1999). *American universities and colleges*. New York, NY: Walter de Gruyter.

Okten, C., & Weisbrod, B. (2000). Determinants of donations in private nonprofit markets. *Journal of Public Economics*, *75*, 255–272.

Posnett, J., & Sandler, T. (1989). Demand for charity donations in private non-profit markets: The case of the U.K. *Journal of Public Economics*, *40*, 187–200.

Schwartz, R. A. (1970). Personal philanthropic contributions. *Journal of Political Economy*, *78*, 1264–1291.

Sebastian, P. (1994, January). Boys town files suit over poor showing in a ratings guide. *Wall Street Journal*, *24*, B6A.

Steinberg, R. (1986a). Should donors care about fundraising? In: S. R. Ackerman (Ed.), *The Economics of Nonprofit Institutions* (pp. 347–364). New York, NY: Oxford University Press.

Steinberg, R. (1986b). The revealed objective functions of nonprofit firms. *Rand Journal of Economics*, *17*(4), 508–526.

The Handbook of Private Schools (1999). Boston, MA: Porter Sargent.

The Official Museum Directory (2000). Washington, DC: National Register Publishing.

Tinkelman, D. (1998). Differences in sensitivity of financial statement users to joint cost allocations: The case of nonprofit organizations. *Journal of Accounting, Auditing and Finance*, *13*(4), 377–394.

Tinkelman, D. (1999). Factors affecting the relation between donations to not-for-profit organizations and an efficiency ratio. *Research in Governmental and Nonprofit Accounting*, *10*, 135–161.

Weisbrod, B. A., & Dominguez, N. D. (1986). Demand for collective goods in private markets: Can fundraising expenditures help overcome free-rider behavior? *Journal of Public Economics*, *30*, 83–96.

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, *48*, 817–838.

A FINANCIAL RATING SYSTEM FOR CHARITABLE NONPROFIT ORGANIZATIONS

John M. Trussel and Janet S. Greenlee

ABSTRACT

We develop a model that uses financial (accounting) indicators to assess financial distress in charitable nonprofit organizations ("charities"), and we use this model to produce a financial rating system for each broad sector of charities: arts, education, human services, public benefit, health and other. The user of this system can assess the likelihood of financial distress, identify the key factors affecting the likelihood of financial distress, and compare the likelihood of financial distress to other charities in its sector.

Our model employs four indicators of financial distress (the equity ratio, the surplus margin, the revenue concentration index and the administrative cost ratio) and controls for both size and sector. Using financial information on charitable organizations provided by the National Center for Charitable Statistics (NCCS) and employing logistical regression, we find that our model of financial distress is statistically significant and fits the data well. Using the results of the logit model, we rank the charities by sector according to their financial indicators and their predicted probabilities.

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 93–116

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11005-7

1. INTRODUCTION

Although several companies, such as Standard and Poor and Moody's, rate the financial condition of for-profit organizations, no financial rating system has been developed for nonprofit organizations. The two types of organizations are similar in at least one important respect: both rely on external sources of capital. The resources available are limited, and resource providers must decide which organizations deserve those resources (Gronbjerg, 1993). In the for-profit domain, one of the factors considered in the allocation of these scarce resources is the organization's financial condition (Franks & Torous, 1989; Jones, 1987; Zavgren, 1983). Our purpose is to develop a system to rate the relative financial condition for one category of nonprofit organizations – charitable nonprofit organizations ("charities"). Our rating system uses financial (accounting) indicators to predict a charity's likelihood of financial distress, to identify the key financial factors that affect this likelihood, and to benchmark this likelihood with other organizations in its sector.

Although previous research on the importance of financial information in the nonprofit sector is predominantly normative (e.g. AAA, 1989; AICPA, 1972; Elkin & Molitor, 1984; Gambino & Reardon, 1981), some empirical research has been conducted (e.g. Sarver, 1994; Steinberg, 1985; Weisbrod & Dominguez, 1986). The focus, however, is on the value of financial information to donors, rather than on the predictive ability of the information. We extend previous empirical research in the charitable nonprofit sector by focusing on the ability of financial information to predict the financial condition of the charity rather than on the impact of such information on decision-makers.

Basic credit information can be readily obtained from such organizations as Dun and Bradstreet. However, a more comprehensive understanding of the relationships between financial indicators, financial distress and financial ratings in the charitable sector is of interest to a variety of groups. These groups include auditors when conducting analytical reviews, government agencies when setting policies and monitoring grants and contracts, managers and board members during the strategic planning process, and potential donors when allocating resources (Greenlee & Trussel, 2000).

We develop and test a model for predicting financial distress in charitable organizations using a methodology that has been employed extensively in the for-profit sector (Altman, 1968; Beaver, 1966; Jones, 1987; Zavgren, 1983) and more recently in the nonprofit sector (Greenlee & Trussel, 2000). This methodology has three basic parameters: (1) accounting data are used as predictor variables of financial distress; (2) the accounting data are based on financial information measured prior to the development of the financial distress (in order

to develop a prediction of financial distress); and (3) the validity of the model is tested on a holdout sample of organizations.

We define a charity as being financially distressed if it has a significant overall decrease in net assets over time. Our model controls for size and sector and uses the four financial indicators of financial distress originally described by Tuckman and Chang (1991): equity ratio, revenue concentration index, surplus margin, and administrative cost ratio. We develop and test our model using financial information from an IRS database developed by the National Center for Charitable Statistics (NCCS).

Our analysis proceeds in four stages. First, for the initial (estimation) sample, we randomly chose approximately one-half of the organizations in the NCCS database to develop a logit model, and we test the significance of this model and the financial indicators. We find the model to be significant.

Second, using the results of the logit model, we compute the probability of financial distress for each charity in the initial sample. The financial indicators for the financially distressed charities are measured prior to their becoming financially distressed. We use these probabilities to predict the future status of each organization as either financially distressed or not financially distressed. We are able to predict the status of the organizations with reasonable accuracy. Third, the model is tested for validity on two alternative samples of organizations. The results on the alternative samples are similar to the initial sample.

Lastly, we create a system to rate the relative financial condition of charities within each sector. The charities are ranked by sector according to their financial indicators and predicted probabilities of financial distress.

The remainder of this paper is organized as follows. In Section 2 we provide an overview of the methods for defining and predicting financial distress in for-profit and charitable organizations. We describe the data and methods in Section 3. Section 4 includes the results of the analyses. The last section presents a summary and a discussion of both the contributions and limitations of this study.

2. BACKGROUND

Two national organizations presently evaluate charities: the Philanthropic Advisory Service of the Better Business Bureau (PAS),[1] and the American Institute of Philanthropy (AIP). Each organization claims part of its mission to be the improvement of donor giving decisions by promoting the public disclosure of financial, governance, and other information.[2] However, none of them provide an overall rating of the financial condition of a charity, and none attempt to predict financial distress.

2.1. Defining a Financially Distressed Organization

Most of the financial distress prediction models in the for-profit sector use financial ratios as predictors of financial distress.[3] The vast majority of the research done in the field of financial distress prediction since the early works of Beaver (1966) and Altman (1968) has been methodological refinements. There are few theoretical refinements of these works, and the lack of a theoretical framework continues to be a problem in the financial distress prediction models. The models have been justified on a more practical basis of predicting financial distress in credit and similar decisions (Jones, 1987; Zavgren, 1983).

Most prediction models define organizations as financially distressed if they file for bankruptcy. Some studies, however, use an alternative construct, since not all financially distressed firms declare bankruptcy. Other avenues may be available for firms in financial distress, such as merger, debt restructure, or voluntary liquidation. Conversely, firms that declare bankruptcy may not be financially distressed. For example, a firm may declare bankruptcy to thwart a labor dispute (Franks & Torous, 1989). Gilbert, Menon, and Schwartz (1990) find that firms selected on the basis of a bankruptcy declaration do not adequately represent the population of all financially distressed firms, since many distressed firms do not declare bankruptcy. They define a financially distressed firm as one that reported cumulative net losses over a three-year period. In their study, they selected samples of financially distressed firms, bankrupt firms, and "healthy" firms (i.e. those that were neither financially distressed nor bankrupt). They find that their prediction model works relatively well in distinguishing between the healthy and the bankrupt firms, but not very well in distinguishing between the financially distressed and the bankrupt firms.

The nonprofit sector differs from the for-profit sector in that, under the federal bankruptcy code, a nonprofit organization cannot be forced into involuntary liquidation or reorganization (11 U.S. C. A. § 303 (a)). A nonprofit organization may, however, dissolve either voluntarily or involuntarily (by state judicial order) for one or more of several reasons, including failure to conduct an activity for a specified period of time (e.g. Cal. Corp. Code § 6610, 8610) or abandonment of the activity of the organization. In addition, a nonprofit organization can dissolve voluntarily for several other reasons, including bankruptcy, disposition of all corporate assets, and loss of all corporation members. A nonprofit organization can be dissolved *involuntarily* for reasons including not having sufficient assets to discharge liabilities, board deadlock, internal dissension among members, fraudulent mismanagement, waste of corporate assets, and failure to adhere to filing or record-keeping requirements (R.M.N.C.A. § 14.20; N.Y. Tax L. § 203-a(1); Fishman & Schwarz, 1995). Hager et al. (1996), in their survey of

NPOs in Minnesota, find that most nonprofits that dissolved do so voluntarily because they were either unable or unwilling to carry out the original activities of the organization.

Because the relationship between charitable organizations and bankruptcy is complex, we do not define a financially distressed charity as one that declared bankruptcy. Rather, we use the broader approach employed by Gilbert, Menon, and Schwartz (1990). As discussed in detail in the next section, we define a financially distressed organization as one with a significant reduction in net assets over time.

2.2. Financial Distress Models for Charities

Although financial distress prediction models are used extensively in the for-profit sector, few such prediction models have been developed for the nonprofit sector. However, Tuckman and Chang (1991) develop a theoretical model that includes four financial indicators of a financially distressed charity. They define a charity as financially distressed if "it is likely to cut back its [program] service offerings immediately when it experiences a financial shock" (p. 445). Such shocks might include an economic downturn or the loss of a major donor. The four hypothesized indicators of financial distress for a charity are inadequate equity balances, high revenue concentration, low administrative costs, and low operating margins.

Tuckman and Chang studied a random sample of 4,730 charitable organizations that filed an annual Form 990 tax return with the IRS in 1983. Charities falling into the lowest quintile for all four variables were defined as being "severely at risk" of becoming financially distressed, and charities with any one of the four variables in the bottom quintile were defined as "at risk" of becoming financially distressed. Tuckman and Chang's empirical tests were descriptive, not predictive, in nature. They made no attempt to determine whether or not those organizations at risk actually became financially distressed.

Greenlee and Trussel (2000) extend Tuckman and Chang's research by using methodologies common in the for-profit sector to develop a model to predict distress for charities. They find a significant relationship between financial distress and three of the four variables adapted from those used by Tuckman and Chang. We expand Greenlee and Trussel's model in three ways. First, we control for both the size and the sector of the organization. Second, we test our model using alternative definitions of financial distress and alternative time periods. Third, we develop a financial rating system based upon the results of the financial distress model.

3. THE FINANCIAL DISTRESS PROBABILITY MODEL

In this section, we describe the data and methodology used to estimate and test a financial distress probability model. The discussion includes the specification of an econometric model, the selection of a set of independent variables, the sample selection, and the estimation methods employed.

3.1. The Financial Distress Indicators

We employ the following probability model to specify the functional relationship between the characteristics of a charity and its probability of being financially distressed in a given period. Let $p(i,t)$ be the probability that charity i will be financially distressed in period t, $x(i,t)$ a vector of attributes of the charity to be measured, and β a vector of unknown parameters to be estimated. Then, $p(i,t)$ is a logit probability function of the measured attributes of the charity (Palepu, 1986):

$$p(i,t) = \frac{1}{1 + e^{-\beta x(i,t)}}.$$

We hypothesize that financial distress in a given period t is predicted by accounting variables in the previous year $t-1$. We use the indicators of financial distress espoused by Tuckman and Chang (1991) to develop our prediction model: equity ratio, revenue concentration, operating margin, and administrative cost ratio. In addition to these four indicators, we control for the size and the sector of the organization.

Equity Ratio (EQUITY). Charities with relatively large amounts of equity are better able to replace revenues following the loss of a major donor or some similar financial problem, than those with relatively small amounts of equity. A charity with a high equity ratio can leverage its assets rather than reduce program services. Thus, the higher the equity ratio, the less likely the organization is to become financially distressed – a predicted negative relationship. This indicator is measured as the ratio of total equity to total revenue.

Revenue Concentration Index (CONCEN). Charities earn revenue from gifts, grants, program services, membership dues, sales of inventories, and investments. Charities with few revenue sources are more vulnerable to financial distress than those with multiple revenue sources. A charity with multiple sources might be able to rely on alternative sources of funding and thus avoid reducing its program services. In other words, charities receiving revenues from fewer sources are more likely to become financially distressed – a predicted positive relationship. Following Tuckman and Chang, the index is constructed similar to the Herfindahl

Index (Weinstock, 1982). The square of the percentage share that each of the organization's revenue sources represents to total revenue is summed. If a charity receives all of its revenue from one source, its revenue concentration index will be one. Conversely, the index of a charity with multiple sources of revenue will approach zero (Weinstock, 1982).

Administrative Cost Ratio (ADMIN). According to Tuckman and Chang, charities with relatively low administrative costs are more vulnerable to financial distress than those with relatively high administrative costs. They hypothesize that an organization with higher administrative costs has more organizational slack and can reduce discretionary administrative costs before reducing its program services. Thus, charities with lower administrative costs are more likely to become financially distressed – a predicted inverse relationship. Following Tuckman and Chang, this variable is measured as the ratio of administrative costs to total expenses.

Surplus Margin (MARGIN). Charities with a relatively low surplus (i.e. excess of revenues over expenses) are more vulnerable to financial distress than those with a relatively high surplus. That is, an organization with a high surplus can operate with a reduced surplus rather than resorting to a reduction in its program services. Thus, the lower the surplus margin, the greater the likelihood of financial distress – a predicted inverse relationship. Following Tuckman and Chang, we measure this variable as the ratio of revenues less expenses to total revenues.

SIZE. Small charities are more vulnerable to financial distress than large ones. Factors such as economies of scale related to costs, age, reputation, and the like are normally correlated with size (Ohlson, 1980; Tinkelman, 1999). Thus, the smaller the charity, the greater the likelihood of financial distress – a predicted inverse relationship. Following Tinkelman (1999), size is measured as the natural log of total assets.

Sector ($SECTOR_j$). Macro-economic factors impact different sectors of charitable organizations differently. For example, some types of charities may have a different reaction to a recession than others. The National Taxonomy Exempt Entities (NTEE) defines 10 categories of charities; however, five of these categories have relatively few organizations in them. Following Tuckman and Chang and Greenlee and Bukovinsky (1998), we merge the organizations in these five categories into one "other" category. Thus, we used five dummy variables to control for the six broad sectors: Arts, Culture and Humanities ($SECTOR_1$); Education ($SECTOR_2$); Human Services ($SECTOR_3$); Public and Societal Benefit ($SECTOR_4$); Health ($SECTOR_5$); and Other. Descriptions of the types of organizations included in each category are presented in Table 1.

The above discussion leads to the identification of six potential determinants of a charity's financial distress probability. The six financial distress indicators and their implied measures are summarized in Table 2. The hypothesized sign of each

Table 1. Examples of the Types of Organizations Included within the Broad Sectors of Charities.

Sector	Examples of Organizations
Arts, culture, humanities	Museums, Galleries, Symphony Orchestras
Education	Nursery Schools, Vocational Schools, Adult Continuing Education Facilities, Libraries
Human services	Social Services, Recreational Services
Public and societal benefit	Public and Citizen Safety, Youth Activities, Public Preparedness
Health	Hospitals, Mental Health Treatment, Rehabilitation Services, Health Support Services, Public Health
Other	Water Conservation, Animal Safety, International Exchanges

variable shows whether the financial distress probability is expected to increase or decrease with that variable.

3.2. Data Collection and Measurement

The data for the study were obtained from the IRS Statistics of Income (SOI) database developed by the National Center for Charitable Statistics (NCCS). This database is limited to nonprofit organizations that are recognized as tax-exempt under the Internal Revenue Code section 501(c)(3) (i.e. charities), are required to file a Form 990 with the IRS, and are not private foundations.[4] The database includes all (approximately 10,000–11,000) charitable organizations with more

Table 2. Measurement of the Variables Used to Indicate Potential Financial Distress in Charities.

Indicator	Measure	Expected Relation with Financial Distress
1. Equity ratio (EQUITY)	Total Equity[a]/Total Revenue	–
2. Revenue concentration (CONCEN)[b]	$\sum(\text{Revenue}_j/\text{Totalrevenues})^2$	+
3. Surplus margin (MARGIN)	Total revenues – Total expenses/Total revenues	–
4. Administrative cost ratio (ADMIN)	Administrative expenses/Total expenses	–
5. Size (SIZE)	Natural log of total assets	–
6. Sector (SECTOR_j)	Dummy variable[c]	?

[a] Equity represents net assets, including restricted and unrestricted assets.
[b] Revenue_j is the revenue from source j.
[c] See Table 1 for description of sectors. Variable is measured as 1 if the organization is a member of sector j and 0 otherwise.

than $10 million in assets and a random sample of approximately 2,000 smaller charitable organizations. Two broad categories of charities are not required to file a Form 990 with the IRS and are therefore not represented in this database. The first category includes all organizations with less than $25,000 in gross receipts ($37,500 in the first three years of operations). The second category is religious organizations, such as churches. The database does not include nonprofit organizations other than charities, such as associations and clubs.

To test the model developed in the previous section, we compare a sample of charities that are financially distressed with a sample of charities that are not financial distressed. To precisely determine when an organization becomes financially distressed is problematic since a decline in actual and reported financial condition likely occurs over time (Gilbert et al., 1990).[5] Gilbert, Menon, and Schwartz (1990) defined a financially distressed for-profit organization as one that had cumulative net losses over a three-year period. Similarly, in a charity, financial distress is assumed to cause a significant reduction in net assets over time, which would manifest through a reduction in revenues and/or an increase in expenses. A nominal reduction in net assets may not be due to financial distress; therefore, for our purposes, a financially distressed charity is defined as an organization that has a significant net reduction in its equity balance (i.e. net assets) over a three-year period. We use two different definitions of a "significant reduction" in net assets as a robustness test. In the first model (model 1), a significant reduction is defined as a 20% reduction in net assets over the three-year period. In the second model (model 2), we require a 50% reduction in net assets over the three-year period.

Specifically, we consider an organization to be financially distressed if net assets at the end of 1995 are significantly less than net assets at the end of 1992. All other organizations are not considered to be financially distressed. In order to develop a predictive model, we measure the variables using data from the 1992 tax year. One problem with this methodology is that a financially distressed charity that became distressed prior to 1992 and did not have a significant decrease in net assets from 1992 to 1995 is classified as not financially distressed. This problem biases the tests against the discrimination of financially distressed charities from those that are not, and makes the tests more conservative.

To be included in the final sample, a charity had to have 1992 data available to compute all of the variables and have net asset balances available for 1995. As noted in Table 3, a large percentage of charities (51.6%) were eliminated from testing, primarily due to not having all of the necessary data available in 1992. We analyzed those charities eliminated and did not find a bias toward the size or sector of the charity. Our final sample consisted of the 6,795 charities. Based upon the classification scheme described above for model 1 (model 2), 718 (504) of the sample organizations are financially distressed and 6,077 (6,291) are not financially distressed. Table 3 summarizes the sample by sector.

Table 3. Composition of a Sample of U.S. Tax-exempt Charities that Filed Form 990 with the Internal Revenue Service for the Years 1992–1995.

	Charitable Organizations		
	Number	%	
Charities in database	14,399	100.0	
Less: Data not available for all years 1992–1995	7,425	51.6	
Outliers[a]	179	1.2	
Final sample	6,795	47.2	
By Sector	Financially Distressed	Not Financially Distressed	Total
Model 1			
Arts, culture & humanities	22	269	291
Education	52	1,437	1,489
Human services	264	1,041	1,305
Public, societal benefit	21	340	361
Health	269	2,545	2,814
Other	90	445	535
Total	718	6,077	6,795
Model 2			
Arts, culture & humanities	12	279	291
Education	19	1,470	1,489
Human services	225	1,080	1,305
Public, societal benefit	15	346	361
Health	180	2,634	2,814
Other	53	482	535
Total	504	6,291	6,795

[a] Outliers are defined as those charities with any financial indicator more than 1.5 hspreads above (below) the 75th (25th)%ile of that indicator. An hspread is the length of the interquartile range. This method of identifying outliers was used since the data are not normally distributed.

4. RESULTS

4.1. Data Description

Panel A of Table 4 presents descriptive statistics of the independent variables by sector. Panel A also includes the results of the Analysis of Variance (ANOVA) to test whether or not the mean value of each variable is the same across sectors. The results indicate that the mean value of each variable is different for at least one

Table 4. Descriptive Statistics for a Sample of 6,795 U.S. Tax-exempt Charities that Filed Form 990 with the Internal Revenue Service for the Years 1992–1995 Partitioned by Sector.

Sector	EQUITY	CONCEN	MARGIN	ADMIN	SIZE
Panel A: Means (Standard Deviations) of the financial distress indicators by sector					
Arts	1.9102	0.4797	0.0892	0.2396	16.2077
	(1.5693)	(0.2082)	(0.1289)	(0.1274)	(2.1853)
Education	1.6514	0.5862	0.0946	0.1832	17.1704
	(1.1486)	(0.1804)	(0.1062)	(0.1195)	(1.6131)
Human services	0.9031	0.6974	0.0450	0.1407	15.7828
	(1.3073)	(0.2327)	(0.1102)	(0.1050)	(1.7804)
Public benefit	1.2382	0.6994	0.0668	0.1497	16.6801
	(1.3959)	(0.2239)	(0.1103)	(0.1034)	(1.6445)
Health	0.6562	0.8661	0.0450	0.1459	17.3670
	(0.7779)	(0.1606)	(0.0746)	(0.1107)	(1.4514)
Other	0.8002	0.7536	0.0338	0.1525	15.5077
	(1.0968)	(0.2316)	(0.1050)	(0.1220)	(2.3637)
All sectors	0.9865	0.7402	0.0545	0.1571	16.6011
	(1.1577)	(0.2287)	(0.1004)	(0.1163)	(1.9683)
F Value	230.26	542.35	84.43	59.81	325.02
(*p*-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel B: Homogeneous subsets by sector					
Arts	5	1	4	3	3
Education	4	2	4	2	5
Human services	2	3	2	1	2
Public benefit	3	3	3	1	4
Health	1	5	2	1	6
Other	2	4	1	1	1

Note: The sample consists of 6,795 charities (see Table 3). The variables EQUITY, CONCEN, MARGIN, ADMIN, and SIZE are defined in Table 2. Panel A presents the mean (standard deviation) for each financial indicator. The hypothesis for each variable is that the means for all of the sectors are equal. Panel B presents the results of Post Hoc tests from Analysis of Variance using the Duncan method. The homogeneous subsets of sectors for the financial indicators of financial distress are displayed. For each variable, there are multiple subsets of sectors, with subset one representing the subset with the lowest values for the variable. For example, the Health sector has the least amount of EQUITY and the Arts sector has the most. Those with multiple sectors in a subset represent sectors with insignificant differences among those sectors in the values of that variable. For example, the Human Services, Public Benefit, Health and Other sectors all have the least amount of ADMIN and are labeled subset 1.

of the sectors. We use Duncan's *Post Hoc* tests to determine which sectors have different means.

Panel B of Table 4 presents the homogeneous subsets of variables from the results of Duncan's *Post Hoc* tests. Inclusion of a sector in a subset indicates that the mean of the variable is not significantly different from the other sectors within that subset. For example, SIZE, with six homogenous subsets, is significantly different for each sector, while CONCEN is not significantly different for the Human Services and Public Benefit sectors (i.e. both sectors are included in subset "3"). The subsets are ranked from the lowest mean (labeled subset 1) to the highest for each variable. For example, Health has the highest CONCEN, while Arts, Culture and Humanities sector has the lowest. In terms of indicating financial distress, the Other sector ranks among the worst in all categories. It has the lowest MARGIN, ADMIN and SIZE and the next to lowest in EQUITY. That sector also has the next to highest CONCEN.

Panel A (Panel B) of Table 5 partitions the descriptive statistics by charities that are financially distressed and those that are not financially distressed for model 1 (model 2). Financially distressed organizations have lower EQUITY, MARGIN and ADMIN ratios, have higher CONCEN indexes and are smaller in SIZE than those organizations that are not financially distressed. The significance of the difference between each of the means of the independent variables is tested using a *t*-test (two-tailed) and a Wilcoxon *z*-test (two-tailed). Each of the differences is significant at the 0.01 level.

Panel C of Table 5 provides a Pearson correlation matrix for the independent variables. All of the variables are significantly correlated at less than the 0.01 level, except SIZE and ADMIN. EQUITY has the highest correlation with both CONCEN (–0.542) and MARGIN (0.436). These correlations with EQUITY appear relatively high and could present a threat to the estimation of the logit model. This problem is considered further in the logit model discussion below.

4.2. Logit Model Estimation

We use a methodology of predicting financial distress in for-profit organizations employed in the literature (e.g. Jones, 1987; Ohlson, 1980; Zavgren, 1983) to test our operationalized financial distress models. Approximately one-half of the organizations in the sample are used to develop logit models, and the remaining organizations are used as holdout samples. The significance of the models is addressed using logit analysis, since the dependent variable is categorical (financially distressed or not) (Maddala, 1991; Palepu, 1986). Using this method,

Table 5. Descriptive Statistics for a Sample of 6,795 U.S. Tax-exempt Charities that Filed Form 990 with the Internal Revenue Service for the Years 1992–1995 Partitioned by Financial Distress Status.

	EQUITY	CONCEN	MARGIN	ADMIN	SIZE
Panel A: Descriptive statistics and univariate tests for model 1					
FD					
Mean	0.2985	0.8177	0.0032	0.1357	16.3688
Std. Dev.	(1.1116)	(0.1983)	(0.1055)	(0.1171)	(2.0046)
NFD					
Mean	1.1205	0.7270	0.0678	0.1610	16.9701
Std. Dev.	(1.1477)	(0.2298)	(0.0959)	(0.1144)	1.6979
t-Statistic	18.208**	−10.131**	18.548**	5.608**	8.822**
Wilcoxon					
z-statistic	18.674**	−11.377**	17.214**	5.505**	7.742**
Panel B: Descriptive statistics and univariate tests for model 2					
FD					
Mean	0.0190	0.7284	0.0166	0.1317	16.2432
Std. Dev.	(1.0609)	(0.2293)	(0.1040)	(0.1145)	(2.0054)
NFD					
Mean	1.1180	0.8397	0.0665	0.1317	16.9594
Std. Dev.	(1.1386)	(0.1883)	(0.0964)	(0.1173)	(1.7089)
t-Statistic	21.678**	−10.613**	18.504**	5.412**	8.930**
Wilcoxon					
z-Statistic	23.022**	−12.547**	17.350**	5.301**	7.794**
Panel C: Pearson correlation coefficients					
CONCEN	−0.542**				
MARGIN	0.436**	−0.310**			
ADMIN	0.159**	−0.215**	0.065**		
SIZE	0.116**	0.065**	0.106**	−0.021	

Note: The sample for model 1 (model 2) consists of 718 (504) charities that are financially distressed (FD) and 6,077 (6,291) charities that are not financially distressed (NFD). The independent variables are defined in Table 2. The SECTOR variables, which are dummy variables for each of the sectors of charities, are not shown.

** Significance at the 0.01 level (two-tailed).

the underlying latent dependent variable, FD^*, is the actual probability of financial distress. This variable is related to the observed variable, FD_i, through the relation:

FD_i = 1 if the organization is financially distressed
FD_i = 0 otherwise

The regression relationship is defined in terms of the latent variable as follows:

$$\mathrm{FD}_i^* = \beta' x_i + u_i,$$

where u_i is an independent and identically distributed random variable, and x_i is the vector of independent (explanatory) variables. The models include all of the independent variables from Table 2. The logit model 1 and model 2 are identical:

$$\mathrm{FD}_i^* = \beta_0 + \beta_1 \mathrm{EQUITY}_{t-1} + \beta_2 \mathrm{CONCEN}_{t-1} + \beta_3 \mathrm{MARGIN}_{t-1} + \beta_4 \mathrm{ADMIN}_{t-1} + \beta_5 \mathrm{SIZE}_{t-1} + \beta_{6-10} \mathrm{SECTOR}_{j,t-1} + u_i$$

The significance of the models overall is addressed using the χ^2 test statistic to test the null hypothesis that all of the parameters in the model were simultaneously equal to zero. The significance of each independent variable is then tested using the *t*-statistic to assess which independent variables in the model significantly impact the probability of financial distress. The results of the tests are included in Table 6.

Overall, both models are significant at the 0.01 level, using χ^2 statistics. Thus, the model fits the data well. For both models, the independent variables EQUITY, MARGIN and SIZE are significant at the 0.01 level (one-tailed test) with the proper signs. The CONCEN and ADMIN variables are not significant at less than the 0.05 level (one-tailed test).[6] Under model 1, three of the five SECTOR dummy variables (representing the Education, Public Benefit, and Health sectors) are significant at less than the 0.05 level (two-tailed test). However, under model 2, only one of the SECTOR dummy variables (representing the Education sector) is significant at less than the 0.05 level (two-tailed test), but two other dummy variables (representing the Public Benefit and the Health sectors) are significant at less than the 0.10 level (two-tailed test).

4.3. Predicting Financial Distress

To test the predictive ability of the financial distress model, we use the logit model developed in the previous section. For example, the observed logistic regression equation for model 1 (from Table 6) is:

$$P(i,t) = \frac{1}{1+e^{-z}}$$

where:

$$Z = 0.5871 - 0.8267\,\mathrm{EQUITY} + 0.2183\,\mathrm{CONCEN} - 5.897\,\mathrm{MARGIN} - 0.2923\,\mathrm{ADMIN} - 0.0915\,\mathrm{SIZE} - 0.4965\,\mathrm{SECTOR}_1 - 1.2597\,\mathrm{SECTOR}_2 - 0.2551\,\mathrm{SECTOR}_3 - 1.1992\,\mathrm{SECTOR}_4 - 0.8338\,\mathrm{SECTOR}_5$$

Table 6. Logit Model of the Probability of Financial Distress for a Sample of 6,795 U.S. Tax-Exempt Charities that Files Form 990 with the Internal Revenue Service for the Years 1992–1995[a].

Variable[b]	Predicted Sign	Model 1		Model 2	
		Coefficient (Std. Error)	*p*-Value[c]	Coefficient (Std. Error)	*p*-Value[c]
Intercept		0.5871	0.3478	0.2063	0.7795
		(0.6253)		(0.7374)	
EQUITY	−	−0.8267	0.0001	−1.7138	0.0001
		(0.1054)		(0.1675)	
CONCEN	+	0.2183	0.5369	0.3898	0.3736
		(0.3536)		(0.4381)	
MARGIN	−	−5.8797	0.0001	−6.0549	0.0001
		(0.8734)		(0.1452)	
ADMIN	−	−0.2923	0.5898	0.0407	0.9502
		(0.5422)		(0.6516)	
SIZE	−	−0.0915	0.0047	−0.1071	0.0047
		(0.0324)		(0.0379)	
$SECTOR_1$	?	−0.4965	0.1946	−0.3236	0.5281
		(0.3828)		(0.5129)	
$SECTOR_2$	?	−1.2597	0.0001	−1.2754	0.0017
		(0.2686)		(0.4072)	
$SECTOR_3$	?	−0.2551	0.2020	0.1623	0.5510
		(0.1999)		(0.2470)	
$SECTOR_4$	?	−1.1992	0.0008	−0.9324	0.0395
		(0.3579)		(0.4529)	
$SECTOR_5$	?	−0.8338	0.0000	−0.4932	0.0401
		(0.1908)		(0.2402)	
Model	χ^{2}[d]	388.456	0.0001	514.523	0.0001

[a] Logit model 1 (model 2) is estimated using a sample of 382 (227) charities that are financially distressed and 3,319 (3,474) that are not financially distressed.

[b] The latent dependent variable (FD) equals one if the charity is financially distressed and zero if the charity is not financially distressed. The independent variables are as follows: EQUITY is the ratio of total equity to total revenue. CONCEN is the sum of the squares of the ratio of each revenue source to the total revenue. MARGIN is the ratio of total revenues less total expenses divided by total revenues. ADMIN is the ratio of administrative expenses to total expenses. SIZE is the natural log of total assets. $SECTOR_j$ is the sector of the charity per Table 1.

[c] All *p*-values on coefficients are based on one-tailed *t*-tests, with the exception of the intercept and SECTOR which are based on a two-tailed *t*-tests.

[d] Model χ^2 is the statistic of a Log-likelihood ratio test.

The predicted dependent variable, the probability of financial distress for organization *i*, is computed using the actual ratios for each organization in the initial sample. The resulting probabilities of financial distress can be used to predict which charities will become financially distressed. Jones (1987) suggests adjusting the cutoff probability for classifying as financially distressed or not financially distressed in two ways. First, we incorporate the prior probability of financial distress, and second, we include the cost of misclassification.

Using logit, the proportion of financially distressed organizations in the sample must be the same as the proportion in the population to account for the prior probability of financial distress. If the proportion is not the same, then the constant must be adjusted (Maddala, 1991). This is more of a problem when a paired sample method is used, which is not the case here. Since we do not know the proportion of financially distressed organizations in the population of all charities, we assume that the proportion in our initial sample is an unbiased estimator of the proportion in the population. Since 10% (7%) of the organizations in the initial sample for model 1 (model 2) are financially distressed, we assume that the prior probability of financial distress is 0.10 (0.07).[7]

The ratio of the cost of type I errors (incorrectly classifying financially distressed organizations as not financially distressed) to type II errors (incorrectly classifying organizations that are not financially distressed as financially distressed) also must be determined. The particular cost function is difficult to ascertain and will depend on the user of the information. For example, a creditor may want to minimize loan losses (and thus type I errors); however, he or she will suffer an opportunity cost (type II error) if the loan is granted to another borrower at a lower rate. In this case, the cost of a type II error is likely to be much smaller that the Type I error rate. Thus, we incorporate several relative cost ratios (and cutoff probabilities) into our analysis. Specifically, we include the relative costs of type I to type II errors of 1:1, 10:1, 20:1, 30:1, 40:1, 60:1, and 100:1.

The results of using the logit models to classify organizations as financially distressed or not are included in Table 7 for the estimation samples and the holdout samples. The cutoff probabilities presented are those that minimize the expected costs of misclassification. Following Beneish (1999), the expected costs of misclassification, ECM, are computed as:

$$\text{ECM} = P(\text{FD})P_{\text{I}}C_{\text{I}} + [1 - P(\text{FD})]P_{\text{II}}C_{\text{II}}, \tag{1}$$

where $P(\text{FD})$ is the prior probability of financial distress, P_{I} and P_{II} are the conditional probabilities of Type I and Type II errors, respectively, and C_{I} and C_{II} are the costs of Type I and type II errors, respectively. Table 7 includes the results for the estimation samples and the holdout samples for model 1 and model 2.

Table 7. Probability Cutoffs that Minimize the Expected Cost of Misclassifying Organizations that are Financially Distressed (Type I Errors) and Organizations that are Not Financially Distressed (Type II Errors).

Cutoff Probability	Costs of Type I to Type II Errors	Type I Error	Type II Error	Overall Error	ECM	ECM Relative to Naïve Strategy
Panel A: Model 1		Estimation	Sample			
0.4500	1:1	0.8438	0.0056	0.0851	0.0894	0.8939
0.1100	10:1	0.3285	0.2568	0.2636	0.5596	0.5596
0.0500	20:1	0.1288	0.5958	0.5515	0.7939	0.8821
0.0050	30:1	0.0161	0.9141	0.8289	0.8710	0.9677
0.0050	40:1	0.0161	0.9141	0.8289	0.8871	0.9856
0.0001	60:1	0.0000	0.9997	0.9048	0.8997	0.9997
0.0001	100:1	0.0000	0.9997	0.9048	0.8997	0.9997
Panel B: Model 1		Holdout	Sample			
0.4500	1:1	0.6871	0.0046	0.0543	0.0514	0.7350
0.1100	10:1	0.2466	0.1377	0.1448	0.2993	0.3219
0.0500	20:1	0.1415	0.3716	0.3541	0.5424	0.5832
0.0050	30:1	0.0567	0.7474	0.6960	0.8117	0.8728
0.0050	40:1	0.0567	0.7474	0.6960	0.8506	0.9146
0.0001	60:1	0.0185	0.9340	0.8659	0.9434	1.0145
0.0001	100:1	0.0185	0.9340	0.8659	0.9934	1.0682
Panel C: Model 2		Estimation	Sample			
0.3500	1:1	0.6012	0.0083	0.0523	0.0498	0.7110
0.1300	10:1	0.2837	0.1045	0.1178	0.2958	0.4225
0.0800	20:1	0.1845	0.2262	0.2231	0.4687	0.3348
0.0800	30:1	0.1845	0.2262	0.2231	0.5979	0.2847
0.0800	40:1	0.1845	0.2262	0.2231	0.7271	0.2597
0.0500	60:1	0.1409	0.3712	0.3541	0.9369	0.2231
0.0100	100:1	0.0615	0.6792	0.6334	1.0622	0.1517
Panel D: Model 2		Holdout	Sample			
0.3500	1:1	0.6528	0.0048	0.0528	0.0371	0.7434
0.1300	10:1	0.3075	0.0908	0.1069	0.2400	0.4800
0.0800	20:1	0.2460	0.1367	0.1448	0.3759	0.3957
0.0800	30:1	0.226	0.1606	0.1654	0.4918	0.5177
0.0800	40:1	0.1845	0.2262	0.2231	0.5839	0.6147
0.0500	60:1	0.1845	0.2262	0.2231	0.7684	0.8089
0.0100	100:1	0.1290	0.4305	0.4081	1.0538	1.1093

Notes: The expected costs of misclassification, ECM, were computed as $\text{ECM} = P(\text{FD})P_I C_I + [1 - P(\text{FD})]P_{II} C_{II}$, where $P(\text{FD})$ is the prior probability of financial distress (0.10 for model 1 and 0.07 for model 2), P_{I} and P_{II} are the conditional probabilities of Type I and Type II errors, respectively. C_{I} and C_{II} are the costs of Type I and Type II errors, respectively.

The results indicate that the models can identify financially distressed organizations; however, the error rates are relatively large in the middle ranges of the relative costs of type I to type II errors. To test the usefulness of the model, we compare these results to a naïve strategy. A naïve strategy would classify all organizations as financially distressed (not financially distressed) when the ratio of relative costs is greater than (less than or equal to) the prior probability of financial distress.[8] If all organizations are classified as financially distressed (not financially distressed), then the naïve strategy makes no type I (type II) errors. In this case, P_{I} (P_{II}) is zero, and P_{II} (P_{I}) is one. Thus, the expected cost of misclassification (1) for the naïve strategy of classifying all charities as not financially distressed (financially distressed) reduces to $0.90C_{\mathrm{II}}$ ($0.10C_{\mathrm{I}}$) for model 1 and $0.93C_{\mathrm{II}}$ ($0.07C_{\mathrm{I}}$) for model 2. We report the ratio of the ECM for our models to the ECM for the naïve strategy in Table 7. For the estimation sample, our models consistently have a lower ECM than the naïve strategy (i.e. the ratio of the relative costs of the model errors to a naïve strategy is less than one). For the holdout sample, our models have a lower ECM than the naïve strategy, except when the ratio of costs is more than 60:1 for model 1 and more than 100:1 for model 2. These results provide evidence to suggest that our financial distress probability model is cost-effective in relation to a naïve strategy for most ranges of the relative costs of type I and type II errors.

4.4. Robustness Tests of the Prediction Models

We test the robustness of the prediction models in two ways. First, we test the models on a holdout sample. The results of the tests on the holdout sample are included in the discussion above. Second, we test the models on a sample of organizations from a different time period. The initial models were developed using data from 1992 to compute the financial indicators, and financially distressed organizations were defined as those with significant declines in net assets from 1992 to 1995. We apply the results of the tests (from Table 6) to the same firms as the initial tests but use data from the 1993 to compute the ratios and define financially distressed organizations as those with significant declines in net assets from 1993 to 1996. The results of these tests are included in Table 8, panels A and B, using the same cutoff probabilities that are used in the estimation sample (from Table 7). The model proves to be cost-effective relative to a naïve strategy, except when the relative cost ratios are 40:1 or greater for model 1 and 60:1 or greater for model 2.

Overall, the estimation results provide evidence of a systematic relationship between the probability of financial distress and certain financial distress indicators. The results are robust to different estimates of the prior probability of financial distress, different specifications of the model, various levels of the relative costs of type I and type II errors, and alternative time periods. However, for relatively

Table 8. Probability Cutoffs that Minimize the Expected Cost of Misclassifying Organizations that are Financially Distressed (Type I Errors) and Organizations that are Not Financially Distressed (Type II Errors): Robustness Tests.

Cutoff Probability	Costs of Type I to Type II Errors	Type I Error	Type II Error	Overall Error	ECM	ECM Relative to Naïve Strategy
Panel A: Model 1						
0.4500	1:1	0.8309	0.0101	0.0880	0.0922	0.1025
0.1100	10:1	0.3221	0.2311	0.2398	0.5301	0.5890
0.0500	20:1	0.1546	0.5671	0.5281	0.8196	0.9106
0.0050	30:1	0.0209	0.9107	0.8265	0.8824	0.9805
0.0050	40:1	0.0209	0.9107	0.8265	0.9034	1.0037
0.0001	60:1	0.0209	0.9107	0.8265	0.9452	1.0502
0.0001	100:1	0.0209	0.9107	0.8265	1.0290	1.1433
Panel B: Model 2						
0.4500	1:1	0.6462	0.0130	0.0570	0.0573	0.0616
0.1100	10:1	0.3121	0.1212	0.1345	0.3312	0.3561
0.0500	20:1	0.1846	0.2580	0.2530	0.4984	0.5359
0.0050	30:1	0.1846	0.2580	0.2530	0.6276	0.6749
0.0050	40:1	0.1846	0.2580	0.2530	0.7569	0.8138
0.0001	60:1	0.1341	0.4247	0.4046	0.9580	1.0302
0.0001	100:1	0.0615	0.7059	0.6613	1.0872	1.1691

Notes: The expected costs of misclassification, ECM, were computed as $\text{ECM} = P(\text{FD})P_{\text{I}}C_{\text{I}} + [1 - P(\text{FD})]P_{\text{II}}C_{\text{II}}$, where $P(\text{FD})$ is the prior probability of financial distress (0.10 for model 1 and 0.07 for model 2), P_{I} and P_{II} are the conditional probabilities of Type I and Type II errors, respectively. C_{I} and C_{II} are the costs of Type I and type II errors, respectively.

large ratios of the costs of type I to type II errors, the model does not prove to be cost-effective relative to a naïve strategy.

4.5. Using the Model for Financial Rating

In this section, we adapt a method used by Tuckman and Chang (1991) to develop a system to rate the relative financial condition of a charity. Under Tuckman and Chang's system, a charity is considered to be "at risk" if it has one of the financial indicators of distress in the bottom quintile and "severely at risk" if the charity is in the bottom quintile for all of the financial indicators. They developed the quintiles for their whole sample and for each of the broad sectors of charities. However, the authors did not develop a composite measure of financial distress similar to our probability of financial distress.

Following Tuckman and Chang (1991), the financial distress indicators are divided into quintiles for each of the six sectors. We use this as a basis of a financial rating system for charities. The system consists of three steps:

Table 9. Computing the Financial Indicators using Hypothetical Data.

Ratio	Example	Ratio Computed	Quintile
EQUITY	$\frac{200{,}000}{400{,}000}$	0.500	3
CONCEN	$\left[\left(\frac{100{,}000}{400{,}000}\right)^2 + \left(\frac{300{,}000}{400{,}000}\right)^2\right]$	0.625	2
ADMIN	$\frac{30{,}000}{300{,}000}$	0.100	2
MARGIN	$\frac{400{,}000 - 300{,}000}{400{,}000}$	0.250	5
SIZE	ln (500,000)	13.1224	1
SECTOR	Human Services		

Notes: The financial indicators are defined in Table 2. Based on the expected signs from Table 2, higher quintiles are associated with lower probability of financial distress for all variables except CONCEN. CONCEN has a direct relationship with financial distress. The hypothetical data are: Revenue from source 1, $100,000; Revenue from source 2, $300,000; Total Revenues, $400,000; Administrative Expenses, $30,000; Total Expenses, $300,000; Equity, $200,000; Total Assets, $500,000.

(1) Compute the probability of financial distress and determine whether or not there is an indication of financial distress.
(2) Rank the probability of financial distress according to the sector quintiles.
(3) Determine what is driving the probability by ranking each of the financial indicators according to sector quintiles.

An example of applying the financial rating system using the hypothetical data included in Table 9 follows. The example assumes that the organization is in the human services sector and applies financial distress model 1.

Step 1: Using model 1 (from Table 6) and the hypothetical data from Table 9, the probability of financial distress is as follows:

$$\textit{Probability of financial distress} = \frac{1}{1 + e^{-Z}},$$

where

$$\begin{aligned} Z = {} & 0.5871 - 0.8267(\text{EQUITY}) + 0.2183(\text{CONCEN}) - 0.2923(\text{ADMIN}) \\ & - 5.8797(\text{MARGIN}) - 0.0915(\text{SIZE}) - 0.4965\,\text{SECTOR}_1 \\ & - 1.2597\,\text{SECTOR}_2 - 0.2551\,\text{SECTOR}_3 - 1.1992\,\text{SECTOR}_4 \\ & - 0.8338\,\text{SECTOR}_5 = 0.5871 - 0.8267(0.500) + 0.2183(0.625) \end{aligned}$$

$$-0.2923(0.100) - 5.8797(0.250) - 0.0915(13.1224) - 0.4965(0)$$
$$-1.2597(0) - 0.2551(1) - 1.1992(0) - 0.8338(0) = -2.6448$$

$$\textit{Probability of financial distress} = \frac{1}{1 + e^{-(-2.6448)}} = 0.0663$$

The probability of financial distress computed using model 1 (0.0663) is more than the financial distress cutoff probabilities for all ratios of the relative costs (from Table 7, Panel A), except for the cost ratios of 1:1 (cutoff of 0.4500) and 10:1 (cutoff of 0.1100). This provides an indication that this charity is financially distressed for most ranges of the relative cost ratios.

Step 2: The probability ranks in the second quintile among those in the human services sector, indicating that approximately 20% (60%) of organizations in this sector have lower (higher) probabilities of financial distress.

Step 3: The ranks of the individual financial distress indicators are listed in Table 9. The MARGIN and the SIZE variables have the largest impact on the probability of financial distress. The MARGIN variable, which is negatively correlated with financial distress, ranks in the top quintile for the sector. This ranking reduces the probability of financial distress. The organization, however, is among the smallest organizations in the human services sector. This ranking increases the probability of financial distress, since SIZE is negatively correlated with financial distress.

This system is merely one possible way of measuring and rating financial distress in the nonprofit sector. Obviously, other financial and nonfinancial information should be considered when evaluating resource allocation and other similar decisions. Financial data are subject to inherent limitations, such as alternative accounting methods and discretionary accounting practices. Also, this rating system does not evaluate the effectiveness or quality of the organization's programs, as it focuses on financial distress.

5. SUMMARY AND CONCLUSION

Although accounting information is used extensively in models predicting financial distress in the for-profit sector, little comparable research has been conducted in the nonprofit sector. Our study extends Tuckman and Chang's (1991) theory of financial distress and expands Greenlee and Trussel's (2000) study to develop and test a financial distress prediction model and a financial rating system for one category of nonprofit organizations – charities. Our rating system uses financial (accounting) indicators to predict a charity's likelihood of financially distress, to

benchmark this likelihood with other organizations in its sector, and to identify the key financial factors that affect this likelihood.

The model assumes that a charity is financially distressed if it has a significant reduction in its net assets over time. We adapt Tuckman and Chang's four indicators of financial distress – equity ratio, revenue concentration, operating margin, and administrative cost ratio – and control for size and sector. A predictive model is developed using the methodology from the for-profit financial distress prediction literature. A logit specification of the model is significant. Also, within certain parameters, we are able to predict with reasonable accuracy whether or not a charity became financially distressed. We also developed a financial rating system based on the predicted probabilities of financial distress and the related sector of the charity.

Our findings also have implications for practice. The model can be used by government agencies when determining contract, grant and monitoring decisions, auditors in developing an audit plan, by potential creditors in determining the credit-worthiness of tax-exempt organizations, and by prospective donors in deciding the amount and timing of contributions.

In developing and testing the model, we have potential specification and measurement problems. We assume that a financially distressed charity is one that has a reduction in equity over a three-year period. Also, in selecting our samples, we use a relatively short time period (1992–1996) and limit our study to charitable organizations. This is primarily due to limitations of the database from which the sample was selected. Since little empirical research has been conducted in this area, many avenues are open. Alternative definitions of financial distress, the incorporation of additional financial variables, and extending the time period of the study may improve the model. The inclusion of nonfinancial information, such as the composition of the board of directors or various accounting methods employed by charities, would also be significant extensions of the study. Finally, studies examining the impact of this model on either actual or simulated decisions would be useful.

NOTES

1. The National Charities Information Bureau merged with PAS in Fall 2001 (Penick, 2000).
2. For more information, see NCIB (http://www.give.org), PAS (http://www.bbb.org/pas), AIP (http://www.charitywatch.org).
3. For more complete surveys of the literature, see Zavgren (1983) and Jones (1987).
4. See Gordon, Greenlee, and Nitterhouse (1999) for a discussion of the usefulness and reliability of this database.

5. As stated in Section 2, we do not define a financially distressed charity as one that has declared bankruptcy, since evidence exists that they cease providing program services prior to dissolving (Hager et al., 1996).

6. The insignificance of the CONCEN and ADMIN variables is surprising, especially since the univariate tests indicate significant differences. The problem may be due in part to the potential multicollinearity in the correlation matrix (see Table 5). We refit the regressions eliminating the variables that we find to be insignificant in the logit model (CONCEN & ADMIN), and again eliminating just the variable common in the two highest correlations (EQUITY). These results did not vary significantly in predictive power and are not reported.

7. We evaluated the sensitivity of the model to other specifications of the prior probability of financial distress by using prior probabilities of 0.01, 0.05 and 0.15. These specifications did not alter the results significantly, and the results are not reported.

8. This switch in strategy between classifying all organizations as not financially distressed to classifying all of them as financially distressed occurs at relative cost ratios of 10:1 (1/0.10) for model 1 and approximately 14:1 (1/0.07) for model 2.

ACKNOWLEDGMENTS

We wish to thank the Aspen Institute's Nonprofit Sector Research Fund for support of this study. An earlier version of this paper was presented at the 1999 annual Association for Research on Nonprofit Organizations and Voluntary Action (ARNOVA) conference, and wish to thank the participants for their helpful comments. We also wish to thank William Baber for his comments.

REFERENCES

Altman, E. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *Journal of Finance*, *3*, 589–609.

American Accounting Association (1989). *Measuring the performance of nonprofit organizations: The state of the art*. Sarasota, FL: American Accounting Association.

American Institute of Certified Public Accountants (1972). *Objectives of financial statements for governmental and not-for-profit organizations*. New York: American Institute of Certified Public Accountants.

Beaver, W. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research* (Supplement), 71–111.

Beneish, M. (1999). The detection of earnings manipulation. *Financial Analysts Journal* (September/October), 24–36.

Elkin, R., & Molitor, M. (1984). *Management indicators in nonprofit organizations: Guidelines to selection and implementation*. New York: Peat Marwick.

Fishman, J., & Schwarz, S. (1995). *Nonprofit organizations: Cases and materials*. Westbury, NY: Foundation Press.

Franks, J., & Torous, W. (1989). An empirical investigation of U.S. firms in reorganization. *Journal of Finance* (July), 747–769.

Gambino, A., & Reardon, T. (1981). *Financial planning and evaluation for the nonprofit organization*. New York: National Association of Accountants.

Gilbert, L., Menon, K., & Schwartz, K. (1990). Predicting bankruptcy for firms in financial distress. *Journal of Business Finance and Accounting* (Spring), 161–171.

Gordon, T., Greenlee, J., & Nitterhouse, D. (1999). Tax-exempt organization financial data: Availability and limitations. *Accounting Horizons* (June), 277–299.

Greenlee, J., & Bukovinsky, D. (1998). Financial ratios for use in the analytical review of charitable organizations. *Ohio CPA Journal* (January–March), 32–38.

Greenlee, J., & Trussel, J. M. (2000). Estimating the financial vulnerability of charitable organizations. *Nonprofit Management and Leadership*, *11*(December), 199–210.

Gronbjerg, K. (1993). *Understanding nonprofit funding*. San Francisco: Jossey-Bass.

Hager, M., Galaskiewicz, J., & Bielefeld, W. (1996). Tales from the grave: Organizations' account of their own demise. *American Behavioral Scientist* (August), 975–994.

Jones, F. (1987). Current techniques in bankruptcy prediction. *Journal of Accounting Literature*, *6*, 131–164.

Maddala, G. (1991). A perspective on the use of limited-dependent and qualitative variables models in accounting research. *Accounting Review* (October), 788–807.

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*(1), 109–131.

Palepu, K. (1986). Predicting takeover targets: A methodological and empirical analysis. *Journal of Accounting and Economics*, *8*, 3–35.

Sarver, P. (1994). Charity efficiency: The public's view. *Nonprofit Times* (April), 1.

Steinberg, R. (1985). Optimal fundraising by nonprofit firms. In: *Giving and Volunteering: New Frontiers of Knowledge* (pp. 411–446). Washington, DC: Independent Sector and United Way of America.

Tinkelman, D. (1999). Factors affecting the relation between donations to not-for-profit organizations and an efficiency ratio. *Research in Government and Nonprofit Accounting*, *10*, 135–161.

Tuckman, H., & Chang, C. (1991). A methodology for measuring the financial vulnerability of charitable nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, *20*(4), 445–460.

Weisbrod, B., & Dominguez, N. (1986). Demand for collective goods in private markets: Can fundraising expenditures overcome free-rider behavior? *Journal of Public Economics*, *30*, 83–96.

Weinstock, D. (1982). Using the Herfindahl index to measure concentration. *Antitrust Bulletin* (Summer), 285–301.

Zavgren, C. (1983). The prediction of corporate failure: The state of the art. *Journal of Accounting Literature*, *2*, 1–38.

THE DEMAND FOR AUDITOR SPECIALIZATION AND ITS EFFECT ON PRIVATE COLLEGE AND UNIVERSITY AUDIT FEES

Mary L. Fischer, Laurence E. Johnson
and Randal J. Elder

ABSTRACT

We study the demand for auditor specialization and the effect of auditor specialization on audit fees for 187 private colleges and universities. We identify two Big 5 firms as specialists based on sample market share. The choice of a specialist auditor is positively related to entity enrollment and endowment assets. We also find that Big 5 specialist auditors receive higher audit fees than non-specialist Big 5 firms. Collectively, the evidence suggests a positive demand for specialist auditors in the private college and university audit market, with returns to specialization in the form of higher market share and higher fees.

1. INTRODUCTION

Audit fees for commercial and governmental entities have been extensively studied. However, comparatively little research has been performed on audit fees of

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 117–133

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11006-9

not-for-profit entities, and we are not aware of previous audit fee research for private colleges and universities.[1] We report the results of a study of private university audit fees for fiscal year 2000. The university environment provides an opportunity to investigate the demand for specialized auditors and the effects of auditor specialization on audit fees. A demand for specialization is plausible in the university market because of the size and complexity of many of the entities in this market.[2]

We identify two Big 5 firms as specialists based on market share. Analysis indicates that the choice of a specialist auditor is positively related to entity enrollment and endowment assets. We find that specialist auditors receive higher fees compared to non-Big 5 firms, and also receive higher fees than non-specialist Big 5 firms. Collectively, the evidence suggests a positive demand for specialist auditors in the private university audit market, with returns to specialization in the form of higher market share and higher fees.

The remainder of the paper is organized as follows. The literature review and research hypotheses are presented in the next section. The third section includes the sample selection and description of the data. The fourth section discusses the empirical test results. The final section provides a summary of the implications and limitations of the study, and suggestions for future research.

2. LITERATURE REVIEW AND DEVELOPMENT OF RESEARCH HYPOTHESES

2.1. Demand for Auditor Specialization

We provide a brief overview of factors influencing the demand for specialized auditors. Wallace (1980) identifies agency costs as one source of audit demand. Early empirical studies focused on the demand for auditing arising from agency costs in non-regulated environments. For example, Chow (1982) finds the demand for auditing by public companies in 1926 is related to agency measures based on size and debt. In the governmental sector, Zimmerman (1977) finds that the demand for auditing is related to the city-manager form of government, an agency measure.

Empirical studies by Francis and Wilson (1988) and DeFond (1992) find that the demand for audit *quality* by public companies, measured by choice of a Big 8 auditor, is related to measures of agency costs. Copley et al. (1995) find that the demand by municipalities for auditor reputation, a surrogate for quality, is related to entity size, as well as risk and leverage measures. They measure auditor reputation using both the Big 8 classification, and a continuous measure that includes auditor size and the number of sample municipalities audited by the

firms. More recently, DeFond et al. (2000) investigate the association between auditor *specialization* and audit fees in the Hong Kong commercial audit market. Their findings indicate that Big 6 specialists are "quality-differentiated auditors that are demanded as an increasing function of agency costs" (p. 62).

We predict a positive demand for specialist auditors in the university audit market related to measures of agency costs. Accordingly, we state the following hypothesis:

H_1. The demand for specialist auditors by private colleges and universities will be positively related to proxies for agency costs.

We test this hypothesis empirically using a logistic regression demand model. Based on research in other settings suggesting that the demand for audit quality is related to size and debt, our demand model includes measures for total student enrollment and long-term debt. Many universities have significant endowment assets. The complexity of auditing endowment agreements and the need to demonstrate accountability to donors may increase the demand for specialized auditors. Similarly, the complexity of compliance auditing for federal government funding may increase the demand for specialist auditors. Our model includes an indicator variable for whether the costs of the A-133 single audit were separately billed as a measure of single audit complexity.

In summary, tests of the demand for auditor industry specialists are based on the following logistic regression model:

$$\text{B5S} = b_0 + b_1\text{LENROLL} + b_2\text{LENDOW} + b_3\text{LLTD} + b_4\text{A133} + \varepsilon \qquad (1)$$

Where:

B5S	= 1 if the auditor is a Big 5 specialist, 0 otherwise
LENROLL	= natural logarithm of student enrollment
LENDOW	= natural logarithm of the total endowment
LLLTD	= natural logarithm of total long-term debt
A133	= 1 if the cost of the A-133 audit is separately billed, 0 otherwise

2.2. Auditor Specialization and Audit Fees

There is an extensive audit fee literature for commercial entities, beginning with Simunic (1980). These studies focus on the structure of the audit market (e.g. Simunic, 1980), auditor reputation and specialization (e.g. Palmrose, 1986) and the effect of non-audit services (e.g. Simunic, 1984). Similarly, many researchers investigate the determinants of audit fees in the governmental sector, beginning

with Rubin (1988), and address issues such as auditor size (Copley, 1989), auditor specialization (Ward et al., 1994) and non-audit services (Elder et al., 1999). These studies also address unique political and agency factors that affect governmental audit fees.

Extant studies of the effect of industry specialization on audit fees report conflicting results. Palmrose (1986) examines the relationship between audit fees and auditor industry specialization for publicly-traded companies in the United States, and finds no significant effect. In contrast, Craswell et al. (1995) observe a Big 5 specialist premium for publicly-traded companies in Australia. DeFond et al. (2000) find that a non-Big 5 specialist charges lower fees in the Hong Kong market for real estate companies.

In the governmental sector, Ward et al. (1994) find that a non-Big 5 specialist in the Michigan governmental market receives higher audit fees. In contrast, Chase (1999) finds that a non-Big 5 specialist charges lower fees for Virginia counties. The conflicting results for auditor industry specialization suggest the need for further research. Studies that find higher fees for specialist auditors often argue that product differentiation by specialist auditors should allow them to charge higher fees. Other studies suggest that specialists should be able to charge lower fees due to economies of scale. It is thus possible that the fee effects of auditor specialization are industry specific.

We suggest that the benefits of auditor specialization (for which auditees should be willing to pay fee premia) should be greatest in specialized industries with complex accounting and auditing requirements. For example, Cullinan (1998) finds that non-Big 5 specialist firms are able to charge higher fees in the audit market for pension plans. The complexities of university accounting and financial reporting suggest that auditor specialization may be important in this industry. Accordingly, we predict that specialist auditors will be able to charge higher fees.

H_2. Audit fees are higher for industry specialist auditors.

Several approaches to measuring auditor industry specialization are possible. Lacking measures of the total university audit market, we measure specialization based on the number of clients audited by each firm in our sample. This is similar to the approach taken in Ward et al. (1994) and Copley et al. (1995). We test this hypothesis with an ordinary least squares regression model of audit fees.

2.3. Control Variables

Additional control variables expected to affect university audit fees included in the audit fee regression model are auditee and auditor size, endowment

assets, long-term debt, single audit requirements, audit delay, and auditor tenure. Consistent with research in the governmental sector that uses population as a measure of size (e.g. Copley, 1989; Rubin, 1988), we use total student enrollment is a measure of university size.[3] Additionally, prior studies document higher audit fees for Big 5 auditors in commercial (e.g. Francis & Simon, 1987; Palmrose, 1986) and governmental environments (e.g. Copley, 1989; Ward et al., 1994). Accordingly, we control for whether the auditor is a Big 5 firm.

As discussed in the previous section, endowment assets and long-term debt are included in the demand for specialist auditors. Larger endowments are expected to be associated with higher fees, since auditing endowment assets and related donor agreements is a significant part of university audits. Several governmental studies use debt variables as measures of client risk (e.g. Chase, 1999; Copley, 1989; Ward et al., 1994).

Private colleges and universities that receive federal funds in excess of $300,000 annually are subject to compliance auditing under the Single Audit Act of 1984 and 1996 amendments (the Act) pursuant to A-133. Several governmental studies find mixed results as to whether single audit requirements are associated with higher audit fees. Baber et al. (1987) report that initial governmental audits under the Act for North Carolina counties are associated with significantly higher fees as compared to fees paid for audit services prior to the Act.

Using more geographically-dispersed data, Raman and Wilson (1992) examine whether implementation of the Act resulted in higher governmental audit fees using a structural change model. They find that implementation of the Act is not associated with higher fees, although internal control weaknesses and situations of non-compliance identified in testing required under the Act had a significant relation with audit fees. Brown and Margavio (1994) found that the presence of single audit requirements has a positive but insignificant effect on audit fees for small Missouri towns. Pearson et al. (1998) find that total accounting and auditing costs fees reported on Form 990 are higher for not-for-profits subject to the Single Audit Act.

Because of their size, most of the universities in our sample are subject to single audit requirements. In approximately 40% of the audits, the audit fee for performance of a single audit was billed separately. There is no theory to explain how components of an audit fee are billed; however, it is plausible that separate billing is an indication of the relative importance of some audit procedures to the overall engagement. If separate billing is an indication that the single audit requirements are particularly significant, then the presence of separate billing may proxy for audit complexity. Accordingly, we include a variable (A133) indicating separate billing of single audit requirements as a proxy for audit complexity in both our audit demand (Eq. (1)) and audit fee (Eq. (2)) models.

Audit delay, measured as the number of days between the balance sheet date and the audit report date, is a proxy for audit complexity and is expected to be positively related to audit fees (Johnson, 1998). Auditor tenure is included to control for fee effects related to the acquisition of audit services, and is expected to have a positive relation with audit fees (Rubin, 1988; Ward et al., 1994). Audit fees may increase with auditor tenure since fees for initial engagements are usually lower (Baber et al., 1987; Chase, 1999) and auditors may be able to increase their fees and receive quasi-rents since it is costly for the client to change auditors (DeAngelo, 1981).

In summary, the following cross-sectional audit fee regression model is the basis for the analysis:

$$\begin{aligned} \text{LFEE} = {} & b_0 + b_1\text{LENROLL} + b_2\text{LENDOW} + b_3\text{LLTD} + b_4\text{A133} \\ & + b_5\text{LDELAY} + b_6\text{LTEN} + b_7\text{B5} + b_8\text{B5S} + \varepsilon \end{aligned} \tag{2}$$

Where:

LFEE = natural logarithm of total audit fees
LENROLL = natural logarithm of student enrollment

Table 1. Summary of Variables and Variable Descriptions for University Specialist Audit Demand and Fee Models.

LFEE	Natural logarithm of audit fees
LENDOW	Natural logarithm of endowment
LENROLL	Natural logarithm of full time equivalent (FTE) student enrollment
LLTD	Natural logarithm of long-term debt
A133	Indicator variable = 1 if audit costs for A133 audit are separately billed; 0 otherwise
LDELAY	Natural logarithm of audit delay (number of days elapsed from balance sheet date to date of the auditor's report)
LTEN	Natural logarithm of auditor tenure in years
B5	Indicator variable = 1 if the auditor is a Big 5 firm, zero otherwise
AA	Indicator variable = 1 if the audit was performed by the former Andersen firm, zero otherwise.
DT	Indicator variable = 1 if the audit was performed by Deloitte & Touche, zero otherwise.
EY	Indicator variable = 1 if the audit was performed by Ernst & Young, zero otherwise.
KPMG	Indicator variable = 1 if the audit was performed by KPMG, zero otherwise.
PwC	Indicator variable = 1 if the audit was performed by PricewaterhouseCoopers, zero otherwise.
B5S	Indicator variable = 1 if the audit was performed either by KPMG or by PricewaterhouseCoopers, zero otherwise.

LENDOW = natural logarithm of the total endowment
LLTD = natural logarithm of total long-term debt
A133 = 1 if cost of the A-133 audit is separately billed, 0 otherwise
LDELAY = natural logarithm of audit delay
LTEN = natural logarithm of auditor tenure in years
B5 = 1 if the auditor is a member of the Big 5, 0 otherwise
B5S = 1 if the Big 5 auditor is an industry specialist, 0 otherwise

Each of the control and research variables is expected to have a positive relation with audit fees. The natural logarithm of the fee and continuous independent variables is used to improve the linear fit of the data, consistent with previous research.[4] Table 1 provides a summary of the variables used in the study.

3. SAMPLE SELECTION AND DATA DESCRIPTION

We sent a survey to 867 college and universities requesting information on audit fees and related variables for fiscal year 2000. Table 2 summarizes the useable responses and the distribution of auditors within the responses.[5] We received a total of 187 useable responses, yielding an overall response rate of about 22%.

Table 2. Summary of Survey Responses to Mail Request to 867 Private Colleges and Universities for Audit Fee and Related Data for Fiscal Year 2000.

Number of responses	
Providing data for variables used in the models	187
Responses representing	
Big 5	
Andersen	7
Deloitte & Touche	12
Ernst & Young	16
KPMG	43
PricewaterhouseCoopers	31
Total	109
Other national	
Grant Thornton	5
McGladrey & Pullen	5
Total	10
Regional and Local	68
Total	187

Table 3. Descriptive Statistics for a National Sample of 187 Private Universities for Fiscal Year 2000 Untransformed Variables Mean (Standard Deviation).

	Big 5 Specialist ($n = 74$)	Big 5 Nonspecialist ($n = 35$)	Total Big 5 ($n = 109$)	Non-Big 5 ($n = 78$)	Total Sample ($N = 187$)
Audit fees ($)	114,989	120,867	116,876	44,459	86,670
	(108,639)	(251,430)	(166,994)	(41,066)	(134,805)
Enrollment	4,947	3,763	4,567	1,887	3,449
(full-time equivalents)	(5,030)	(3,226)	(4,548)	(1,605)	(3,851)
Audit delay in days	82 (36)	84 (25)	83 (33)	75 (39)	80 (36)
Auditor tenure in years	14 (15)	17 (16)	15 (15)	12 (13)	14 (14)
Endowment ($)	278,300,000	554,400,000	366,900,000	36,590,000	229,100,000
	(472,400,000)	(1,642,000,000)	(1,007,000,000)	(47,900,000)	(785,800,000)
Long-term Debt ($)	65,260,000	40,710,000	57,380,000	14,510,000	39,500,000
	(78,820,000)	(62,460,000)	(74,570,000)	(16,880,000)	(61,610,000)
Separate bill for	0.5270	0.5143	0.5229	0.2692	0.4171
A-133 costs	(0.5027)	(0.5071)	(0.5018)	(0.4464)	(0.4994)

We note a concentration of Big 5 auditors in our data and a further concentration of clients within the Big 5. Approximately 60% of the institutions in our sample are audited by the Big 5. KPMG and PricewaterhouseCoopers (PwC) audit 43 and 31 sample firms, respectively, and have significantly more clients than the other Big 5 firms in the sample. Accordingly, based on the distribution of clients among the Big 5 firms in our sample, we classify KPMG and PwC as specialists. Several non-Big 5 national and regional firms have more than one client in the sample, but none has more than seven clients.

Descriptive statistics for the sample are reported in Table 3. The average enrollment of the institutions in our sample is 3,449 full time equivalent students (FTEs), somewhat smaller than the average size of private universities, overall.[6]

The average audit fee for the entire sample is $86,670. The average audit fee of $116,876 for universities audited by Big 5 firms is substantially higher than the average of $44,459 for institutions audited by non-Big 5 firms. Measured by enrollment, institutions audited by Big 5 firms are over twice as large as those audited by non-Big 5 firms. Also, Big 5 auditees have substantially larger endowments than institutions audited by non-Big 5 firms, and are more likely to be billed separately for the costs associated with A-133 audit requirements.

Differences between Big 5 specialists and Big 5 non-specialist firms are less pronounced. The average audit fee for Big 5 specialists of $114,989 is lower than the average fee of $120,867 for non-specialist Big 5 auditors. Universities audited by Big 5 specialists have larger enrollments and more long-term debt, although they have smaller average endowments.

4. RESULTS

Results for the logistic Big 5 specialist audit demand model are reported in Table 4. Choice of a specialist auditor is significantly positively related to enrollment, endowment, and long-term debt. The A133 variable has a positive coefficient, but is not significant at conventional levels. The concentration of auditors in our sample suggests that auditor specialization is important in the university audit market, and the model is consistent with a positive demand for auditor specialization related to size proxies for agency costs.

Results for the primary audit fee regression model are reported in Table 5. White's test (White, 1980) indicates the presence of heteroscedasticity in the data, so we compute standard errors, *t*-stastistics, and *p*-values using White's correction for heteroscedasticity. Wilk-Shapiro statistics indicate that the residuals are approximately normally distributed for each model, and residual diagnostics indicate that results are not significantly influenced by extreme observations.

Table 4. Logistic Regression Model of Demand for Big 5 Specialist Auditors for a National Sample of Private Colleges and Universities for Fiscal Year 2000.

Model : $B5S = b_0 + b_1 LENROLL + b_2 LENDOW + b_3 LLTD + b_4 A133 + \varepsilon$

Variable	Expected Sign	Coefficient	Standard Coefficient/ Error	Standard Error	*p*-Value
Constant	+/−	−17.2857	2.9879	−5.79	<0.001
LENDOW	+	0.2251	0.1192	1.89	0.059
LENROLL	+	0.4639	0.2332	1.99	0.046
LLTD	+	0.5437	0.1551	3.50	<0.001
A133	+	0.2776	0.3573	0.78	0.437
Deviance	200.35				
p-value	0.1671				
Degrees of Freedom	182.0				
Wilk-Shapiro statistic	0.9527				
Proportion of Big 5 Specialist = 0 correctly classified = 0.832					
Proportion of Big 5 Specialist = 1 correctly classified = 0.527					
Overall proportion correctly classified = 0.711					

Note: See Table 1 for description of variables.
$n = 187$.

In addition, no variance inflation factor (VIF) exceeds 2.3 with most less that 2.0, suggesting that coefficient estimates are not unduly affected by collinearity. Accordingly, the regression results can be interpreted in a straightforward manner.

Except for the long-term debt variable, all of the variables are significant and positively related to audit fees.[7] The Big 5 indicator variable is positive and significant, consistent with the extensive body of literature that indicates Big 5 firms receive higher fees. Consistent with the second research hypothesis, audit fees are higher for Big 5 specialist auditors. We report additional fee models using individual audit firm indicator variables and limiting the sample to observations with Big 5 auditors as further specification tests.

As expected, the student enrollment size measure is positive and highly significant. Audit fees are also positively related to the size of the endowment, consistent with greater audit effort for entities with large endowments. Total audit fees are also higher when A-133 audit amounts are billed separately, consistent with greater audit complexity related to single audit requirements.[8] The delay variable is significant, indicating that audit complexity affects audit fees. Auditor tenure is also significant for the full sample, consistent with the argument that costly contracting allows auditors to charge quasi-rents (DeAngelo, 1981).

In Table 6, we replace the Big 5 and Big 5 specialist variables with indicator variables for each Big 5 firm. These results provide further support for the

Table 5. Regression Model of Audit Fee for a National Sample of Private Colleges and Universities for Fiscal Year 2000.

Model : $\text{LFEE} = b_0 + b_1\text{LENROLL} + b_2\text{LENDOW} + b_3\text{LLTD} + b_4\text{A133} + b_5\text{LDELAY} + b_6\text{LTEN} + b_7\text{B5} + b_8\text{B5S} + \varepsilon$

Variable	Expected Sign	Coefficient	Standard Error	*t*-Statistic	*p*-Value[a]
Constant	+/−	5.2403	0.6643	7.8876	<0.001
LENROLL	+	0.3546	0.0821	4.3180	<0.001
LENDOW	+	0.0821	0.0402	2.0398	0.021
LLTD	+	−0.0197	0.0443	−0.4448	0.328
A133	+	0.0945	0.0639	1.4775	0.070
LDELAY	+	0.3500	0.0925	3.7830	<0.001
LTEN	+	0.0654	0.0342	1.9077	0.029
B5	+	0.2581	0.0852	3.0275	0.001
B5S	+	0.1560	0.0946	1.6485	0.050
Adjusted R^2	0.5740				
Model *F*	32.3382				
Prob (*F*)	0.0000				
Wilk-Shapiro statistic	0.9270				

Note: See Table 1 for description of variables. Standard errors, *t*-statistics, and *p*-values are computed using White's (1980) correction for heteroscedasticity.

$n = 187$.

[a] One-tailed test where sign of coefficient is specified, otherwise, two-tailed test.

existence of a specialist fee premium. The coefficient for each firm is positive and at least marginally significant, consistent with a Big 5 premium. Although the coefficient for non-specialist AA (0.4963) is nominally the largest among Big 5 firms, specialists PwC (0.4754) and KPMG (0.3747) are second and third largest and have larger *t*-statistics than the other firms.

As an additional test, we reperform the previous analyses restricting the sample to observations with Big 5 auditors. In Table 7, the coefficient for the specialist indicator variable is positive and significant, indicating the existence of a within-Big 5 specialist premium. Consistent with the results reported in Table 5, the control variables are also positive and significant, with the exception of the long-term debt variable.

We report in Table 8 the results for the Big 5 sample with individual firm indicator variables for AA, DT, KPMG and PwC. We exclude EY to prevent overidentification of the model, since EY had the smallest coefficient of the Big 5 in Table 6. Only the KPMG and PwC indicator variables are significant, consistent with higher fees for Big 5 specialists.

Table 6. Regression Model of Audit Fee with Audit Firm Indicator Variables for a National Sample of Private Colleges and Universities for Fiscal Year 2000.

Model : $\text{LFEE} = b_0 + b_1\text{LENROLL} + b_2\text{LENDOW} + b_3\text{LLTD} + b_4\text{A133} + b_5\text{LDELAY} + b_6\text{LTEN} + b_7\text{AA} + b_8\text{DT} + b_9\text{EY} + b_{10}\text{KPMG} + b_{11}\text{PwC} + \varepsilon$

Variable	Expected Sign	Coefficient	Standard Error	*t*-Statistic	*p*-Value[a]
Constant	+/−	5.3467	0.6585	8.1194	<0.001
LENROLL	+	0.3451	0.0821	4.2037	<0.000
LENDOW	+	0.0799	0.0392	2.0383	0.022
LLTD	+	−0.0208	0.0438	−0.4739	0.318
A133	+	0.0971	0.0697	1.3944	0.082
LDELAY	+	0.3512	0.0918	3.8245	<0.001
LTEN	+	0.0703	0.0346	2.0268	0.022
AA	+	0.4963	0.2701	1.8373	0.034
DT	+	0.2845	0.1413	2.0127	0.023
EY	+	0.1519	0.1036	1.4651	0.072
KPMG	+	0.3747	0.0961	3.8991	<0.001
PwC	+	0.4954	0.1176	4.2097	<0.001
Adjusted R^2	0.5756				
Model *F*	23.9335				
Prob (*F*)	0.0000				
Wilk-Shapiro statistic	0.9304				

Note: See Table 1 for description of variables. Standard errors, *t*-statistics, and *p*-values are computed using White's (1980) correction for heteroscedasticity.

$n = 187$.

[a] One-tailed test where sign of coefficient is specified, otherwise, two-tailed test.

The concentration of clients with KPMG and PwC suggests that there is a positive demand for specialization, with returns to the specialist firms in the form of higher market share. The effect on fees of within Big 5 specialization is consistent with the results in Craswell et al. (1995). That is, there are returns to specialization, and they constitute a significant portion of the Big 5 fee premium.

For non-Big 5 firms, fewer clients may be required to achieve within firm specialization. We classify non-Big 5 firms as specialists if the have more than one client in the sample. Neither the number of clients or a dichotomous specialization measure is significant in any model. Although Ward et al. (1994) and Cullinan (1998) provide evidence on the existence of premium non-Big 5 specialists, our data do not reveal any such specialists in the university audit market.

Table 7. Regression Model of Audit Fee Using Only Big 5 Observations for a National Sample of Private Colleges and Universities for Fiscal Year 2000.

Model : $\text{LFEE} = b_0 + b_1\text{LENROLL} + b_2\text{LENDOW} + b_3\text{LLTD} + b_4\text{A133} + b_5\text{LDELAY} + b_6\text{LTEN} + b_7\text{B5S} + \varepsilon$

Variable	Expected Sign	Coefficient	Standard Error	*t*-Statistic	*p*-Value[a]
Constant	+/−	5.0921	0.7953	6.4024	<0.001
LENROLL	+	0.3316	0.0922	3.5975	<0.001
LENDOW	+	0.1501	0.0456	3.2904	<0.001
LLTD	+	−0.0579	0.0530	−1.0926	0.140
A133	+	0.1238	0.0839	1.4761	0.071
LDELAY	+	0.3192	0.1170	2.7276	0.003
LTEN	+	0.1083	0.0528	2.0496	0.022
B5S	+	0.1787	0.0936	1.9097	0.029
Adjusted R^2	0.4862				
Model F	15.6000				
Prob (*F*)	0.0000				
Wilk-Shapiro statistic	0.9306				

Note: See Table 1 for description of variables. Standard errors, *t*-statistics, and *p*-values are computed using White's (1980) correction for heteroscedasticity.

$n = 109$.

[a] One-tailed test where sign of coefficient is specified, otherwise, two-tailed test.

5. SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FUTURE RESEARCH

We analyze audit fees for private colleges and universities to examine the demand for auditor specialization and its effect on audit fees. This market is interesting because it has not previously been studied, it involves fairly complex accounting and auditing, and it includes entities that vary considerably in size. As such, the university audit market offers a rich opportunity to study the effects of auditor specialization.

Audit fees are higher for universities audited by Big 5 firms. In addition, we find evidence of demand for specialists within the Big 5, and these specialists firms receive a fee premium. This adds to a growing body of research that finds evidence of fee premiums for Big 5 firms (Craswell et al., 1995), and non-Big 5 firms (Cullinan, 1998; Ward et al., 1994).

Our study is based on a study of private colleges and universities, and the results may not extend to public universities, or other not-for-profit organizations.

Table 8. Regression Model of Audit Fee Using Only Big 5 Observations with Audit Firm Indicator Variables for a National Sample of Private Colleges and Universities for Fiscal Year 2000.

Model : $\text{LFEE} = b_0 + b_1\text{LENROLL} + b_2\text{LENDOW} + b_3\text{LLTD} + b_4\text{A133} + b_5\text{LDELAY} + b_6\text{LTEN} + b_7\text{AA} + b_8\text{DT} + b_9\text{KPMG} + b_{10}\text{PwC} + \varepsilon$

Variable	Expected Sign	Coefficient	Standard Error	*t*-Statistic	*p*-Value[a]
Constant	+/−	5.1386	0.8072	6.3653	<0.001
LENROLL	+	0.3177	0.0924	3.4379	<0.001
LENDOW	+	0.1468	0.0425	3.4525	<0.001
LLTD	+	−0.0599	0.0512	−1.1699	0.122
A133	+	0.1256	0.0928	1.3539	0.089
LDELAY	+	0.3208	0.1167	2.7477	0.003
LTEN	+	0.1185	0.0532	2.2279	0.014
AA	+/−	0.3657	0.2768	1.3208	0.190
DT	+/−	0.1545	0.1591	0.9709	0.334
KPMG	+	0.2628	0.1140	2.3055	0.012
PwC	+	0.3752	0.1361	2.7551	0.003
Adjusted R^2	0.4909				
Model *F*	11.4172				
Prob (*F*)	0.0000				
Wilk-Shapiro statistic	0.9432				

Note: See Table 1 for description of variables. Standard errors, *t*-statistics, and *p*-values are computed using White's (1980) correction for heteroscedasticity. The indicator for Ernst & Young (EY) is omitted to prevent model overidentification.

$n = 109$.

[a] One-tailed test where sign of coefficient is specified, otherwise, two-tailed test.

The results are also based on survey responses and are subject to the limitations of survey research, such as the potential for non-response bias. Additional research in other markets is necessary to understand factors that result in a demand for auditor industry specialization, and the circumstances in which specialization results in higher audit fees. Finally, future research should investigate whether specialist auditors are associated with greater audit effort or other measures of audit quality.

NOTES

1. For convenience, we refer to these private colleges and universities as "universities" throughout the text.

2. Private colleges and universities are not-for-profit entities, and accounting rules for these entities are established by the Financial Accounting Standards Board (FASB). Public colleges and universities are governmental entities, and are subject to standards established by the Government Accounting Standards Board (GASB).

3. We also test total revenues as a measure of auditee size. The explanatory power of the model is significantly lower when we use total revenues. We speculate that the enrollment measure, in addition to proxying for size, captures additional services associated with the number of students. Use of total revenues, rather than total enrollment, does not significantly impact the other variables in the model.

4. In many previous studies (e.g. Rubin, 1988; Ward et al., 1994) tenure is not logged. Reported results are not sensitive to use of logged vs. an untransformed tenure variable. We also test a dichotomous variable using a tenure cutoff of 10 years and obtain similar results.

5. The membership list of the National Association of College and University Business Officers was used to identify four-year, degree granting, private not-for-profit colleges and universities in the United States. Member institutions located in Ohio were compared to those listed in the 2000 Higher Education Directory to ensure the sample's validity. The comparison resulted in finding one member institution that was not listed in the Directory. A personally signed letter that explained the information request together with a postage paid envelope was sent to each of the institutions in the sample population.

6. According to the National Center for Education Statistics Mini-Digest (2000), private U.S. colleges and universities have an average enrollment of 4,000 and an average endowment of approximately $700 million.

7. Many prior audit fee studies control for the date of the auditees' year-end. In a preliminary specification of the model, we included a categorical variable for institutions with fiscal years ending on June 30 (the predominant fiscal year end in our sample). We did not predict the sign of the variable because June 30 university audits could reflect either a "slow season" discount relative to the traditional private-sector winter busy season or a premium reflecting a "mini-busy season" for university auditors because of the concentration of work at June 30. The coefficient for this variable was not significant and was excluded from further analysis.

8. Alternative measures based on the presence of a single audit and the number of major programs are not significant. As previously noted, most universities are subject to single audit requirements, and lack of data on the number of programs results in loss of observations. We carefully worded the fee survey to insure that single audit costs were included in the total audit fee. Results for other variables are not significantly affected by inclusion of the A133 variable.

ACKNOWLEDGMENTS

We thank the editor and an anonymous reviewer for their helpful comments. We also thank John Engstrom, Linda Kidwell, Kathryn Jervis, and participants at the 2002 American Accounting Association Annual Meeting and 2002 Government and Non-profit Section Midyear Meeting for helpful comments.

REFERENCES

Baber, W., Brooks, E., & Ricks, W. (1987). An empirical investigation of the market for audit services in the public sector. *Journal of Accounting Research*, *25*(Autumn), 293–305.

Brown, K., & Margavio, T. (1994). Audit costs of small cities in an unregulated audit market environment. *Public Budgeting and Financial Management*, *6*(3), 376–421.

Chase, B. (1999). The influence of auditor change and type on audit fees for municipalities. *Research in Governmental and Nonprofit Accounting*, *10*, 49–63.

Chow, C. (1982). The demand for external auditing: Size, debt and ownership influences. *The Accounting Review*, *57*(April), 272–290.

Copley, P. (1989). The determinants of local government audit fees: Additional evidence. *Research in Governmental and Nonprofit Accounting*, *5*, 3–23.

Copley, P., Gaver, J., & Gaver, K. (1995). Simultaneous estimation of the demand for differentiated audits: Evidence from the municipal audit market. *Journal of Accounting Research*, *33*(Spring), 137–155.

Craswell, A., Francis, J., & Taylor, S. (1995). Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics*, *20*, 297–322.

Cullinan, C. (1998). Evidence of non-big 6 market specialization and pricing power in a niche assurance service market. *Auditing: A Journal of Practice and Theory* (Supplement), 47–57.

DeAngelo, L. (1981). Auditor independence, low-balling, and disclosure regulation. *Journal of Accounting and Economics*, *3*(August), 113–127.

DeFond, M. (1992). The association between changes in client firm agency costs and auditor switching. *Auditing: A Journal of Practice and Theory*, *11*(Spring), 16–31.

DeFond, M., Francis, J., & Wong, T. (2000). Auditor industry specialization and market segmentation: Evidence from Hong Kong. *Auditing: A Journal of Practice and Theory* (Spring), 49–66.

Elder, R., Kattelus, S., & Douthett, E. (1999). Audit fees and nonaudit fees in the governmental sector: A self-selection analysis. *Research in Governmental and Nonprofit Accounting*, *10*, 65–86.

Francis, J., & Simon, D. (1987). A test of audit pricing in the small-client segment of the U.S. audit market. *The Accounting Review*, *62*(January), 145–157.

Francis, J., & Wilson, E. (1988). Auditor changes: A joint test of theories relating to agency costs and auditor differentiation. *The Accounting Review*, *63*(October), 663–682.

Higher Education Directory (2000). Higher Education Publications, Inc., Falls Church, VA.

Johnson, L. (1998). Further evidence on the determinants of municipal audit delay. *Journal of Public Budgeting and Financial Management*, *10*(Fall), 375–397.

Mini-Digest of Educational Statistics (2000). National Center for Educational Statistics. Washington, DC: U.S. Department of Education.

Palmrose, Z. (1986). Audit fees and auditor size: Further evidence. *Journal of Accounting Research*, *24*(Spring), 97–110.

Pearson, T., Brooks, R., & Neidermeyer, A. (1998). The determinants of monitoring costs in not-for-profit organizations. *Public Budgeting, Accounting & Financial Management*, *10*(Winter), 499–512.

Raman, K., & Wilson, E. (1992). An empirical investigation of the market for "single audit" services. *Journal of Accounting and Public Policy*, *11*(Winter), 271–295.

Rubin, M. (1988). Municipal audit fee determinants. *The Accounting Review*, *63*(April), 219–236.

Simunic, D. (1980). The pricing of audit services: Theory and evidence. *Journal of Accounting Research*, *18*(Spring), 161–190.

Simunic, D. (1984). Auditing, consulting and auditor independence. *Journal of Accounting Research*, *22*(Autumn), 679–702.

Wallace, W. (1980). *The economic role of the audit in free and regulated markets*. New York: Touche Ross & Co.

Ward, D., Elder, R., & Kattelus, S. (1994). Further evidence on the determinants of municipal audit fees. *The Accounting Review*, *69*(April), 399–411.

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, *48*(May), 817–838.

Zimmerman, J. (1977). The municipal accounting maze: An analysis of political incentives. *Journal of Accounting Research*, *18*(Supplement), 107–144.

ENDOGENEITY ISSUES IN GOVERNMENTAL RESEARCH

Edward B. Douthett Jr. and Linda M. Parsons

ABSTRACT

We examine the issue of endogeneity in government accounting research. Estimation bias and inconsistency can arise when single equation estimation includes an explanatory endogenous variable. We discuss the modeling and estimation issues involving endogenous explanatory variables and we illustrate alternative specification tests for explanatory endogeneity when a system of equations includes a qualitative dependent variable, a common feature in government research. Our specification testing confirms the presence of endogeneity bias in a model used to estimate auditor reputation and audit fees in the government sector. We conclude the paper by offering suggestions for research where examining or considering explanatory endogeneity using simultaneous equations may be appropriate in other government settings.

1. INTRODUCTION

Accounting studies examining issues in both the for-profit and government arenas focus on a number of methodological problems, including multicollinearity of predictor variables, non-normality of data and heterogeneity of the variance in error terms. A problem too often ignored is the issue of endogeneity of predictor

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 135–149

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11007-0

variables in single equation analysis. We review how the issue of endogeneity has been addressed in previous accounting research, and examine various statistical procedures that could be used in government accounting research when the predictor variables are both endogenously determined and limited or qualitative in their range.

The importance of identifying and correcting for right-hand-side endogeneity is straightforward: inferences from statistical analyses are only as good as the measurement of the constructs and the validity of statistical methods. Although accounting researchers typically take steps to avoid or correct methodological problems, the problem of endogeneity demands rigorous specification testing, relatively cumbersome estimation methods, and most importantly, a theoretical argument as to why the endogeneity exists. Thus, the cost of addressing the endogenous effects may be perceived to outweigh the benefits. However, the consequences of ignoring explanatory endogeneity can be severe, rendering results which are biased and inconsistent (Fomby, Hill & Johnson, 1984). Specification error, such as that resulting from explanatory endogeneity, is particularly problematic in models that use limited dependent variables, including popularly used models such as probit, logit, or tobit (Yatchew & Griliches, 1984).

The remainder of the paper is organized as follows. Section 2 presents specific instances where accounting researchers have successfully employed simultaneous estimation techniques to control for endogeneity. Section 3 includes a discussion of the modeling aspects of simultaneous equations. Section 4 compares alternative testing methods to identify simultaneity bias, and Section 5 offers research suggestions where studying the issue of endogeneity in a government setting may be appropriate.

2. LITERATURE REVIEW

An illustration of endogeneity in the accounting realm is provided by Demski, Fellingham, Ijiri and Sunder (2002). Accounting researchers assume that transactions occur and then accountants measure and report them. However, as Demski et al. (2002) point out, sometimes transactions are designed and executed in order to achieve certain accounting measurement and disclosure. In other words, there are times when the occurrence of a transaction and the related accounting measurement are determined jointly and simultaneously. An example of the estimation problems related to explanatory endogeneity is identified by Leftwich, Watts and Zimmerman (1981). They note that their inability to find a significant relationship between the voluntary provision of interim reports and

certain accounting measures is likely due, in part, to the endogeneity of accounting choice and accounting measurement.

A number of private sector accounting studies utilize simultaneous equations to address the issue of endogeneity. For example, Zeibart and Reiter (1992) theorize that the effective interest rate of a corporate bond (a continuous variable) is determined, in part, by assigned bond ratings (a categorical variable measured on a scale of 1–6) and certain accounting information. However, bond ratings are determined based on a number of factors, including accounting information. Thus, accounting measures affect yield rates directly and indirectly through their impact on bond ratings.[1]

If endogeneity exists, results generated from a set of simultaneous equations can differ dramatically from those obtained from a single regression equation (i.e. where regression assumptions are violated and bias exists). For example, Ireland and Lennox (2002) assess the audit fee premium associated with the Big 6 firms and find an audit fee premium of 55% using a model of simultaneous equations versus a 19% premium using a single equation model. Further, Copley and Douthett (2002) find that, compared to a single equation model, the signs on the coefficients of exogenous variables change when they introduce a full set of simultaneous equations.

2.1. Governmental Entities

Accounting researchers studying questions in the governmental sector also confront the problem of endogeneity. Examining a sample of governmental units that receive Federal funding, Copley, Doucet and Gaver (1994) use Amemiya's (1978) two-stage generalized least squares (2SGLS) procedures to demonstrate that audit fees and audit quality are determined concurrently.[2]

Two additional studies investigate audit issues for municipalities while specifically controlling for endogeneity. Copley, Gaver and Gaver (1995) confirm that the supply of audit services for a certain fee and the demand for auditor reputation are decided in tandem. Elder, Kattelus and Douthett (1999) find that municipalities' purchases of non-audit services and the fees charged for audit services are positively associated and jointly determined.

Studies examining the endogeneity issue for government entities have been limited to auditor choice and fee models. There are likely many other opportunities to study endogeneity in the government setting. In the next section we review the econometric concerns in modeling and estimating simultaneous equations.

3. MODELING AND ESTIMATION OF SIMULTANEOUS EQUATIONS

3.1. The Econometric Issues

A potential endogeneity problem exists when a predictor variable is not determined independently of the dependent variable. Single equation models assume one-way causality. Variables on the left side of the equation are presumed to be determined by those on the right side. The case of simultaneity suggests two-way causality. Variables on the left and right side of the equation are determined jointly.

Estimations where the predictor variable is not corrected for endogenous effects produce error terms that are not independent of the explanatory variables. This is a violation of the classical ordinary least squares (OLS) assumptions and results in biased and inconsistent coefficient estimates that cannot be reliably interpreted (Berry, 1984). Uncorrected right-hand-side endogeneity not only affects the estimation of the coefficient in question, but also the other right-hand-side coefficients in the model, amplifying bias and inconsistency across all parameter estimates in the equation. The use of simultaneous equations models allows accounting researchers to resolve problems of endogeneity by purging the endogenous effects from the explanatory variables, thereby removing any contemporaneous correlation between the explanatory variables and the error term, providing consistent estimates valid for inferences.

3.2. Simultaneous Estimation with Linear Equations

Estimation of simultaneous equations first requires a hypothetical specification of the structural form of the system of equations (we limit our discussion to estimation using two-stage least squares (2SLS)). The structural form equations portray the behavior of an economic agent, and all explanatory variables in the structural equations should be theoretically linked to the dependent variable.

Once the structural form is determined, the reduced form equations are used to obtain instruments, or proxies, for the right-hand-side endogenous variables. Reduced form equations are simply re-expressions of the dependent variables solely in terms of all exogenous variables in the system. Procedurally, the predicted value for each dependent variable is obtained from the reduced form regressions. This predicted value, defined as an instrumental variable, is highly correlated with the original dependent variable and not correlated with error term. Loosely speaking, the interdependencies among the endogenous variables are eliminated, and therefore the predicted values meet the requirements of independence for use

as a right-hand-side variable in ordinary least squares estimation. In summary, the first stage of 2SLS estimation is a regression of the reduced form to obtain useable instruments (i.e. predicted values) in place of the endogenous right-hand-side variables. The second stage of the estimation is the regression of the structural form, including the instruments obtained from the first stage.

3.3. The Importance of the Reduced Form

The primary purpose of the reduced form is to obtain an orthogonal instrument for use in estimating the structural form. However, it is worth noting that the structural and reduced form versions of a simultaneous equations model are equally valid formulations of the same economic process. The two versions introduce different issues for estimation and are suited for different purposes. Specifically, if the primary purpose of the analysis is to predict the dependent variable or determine if a factor is a significant determinant without regard to sign, then after constructing a theoretically valid, unestimated structural model, the practical work of estimation can be conducted entirely in terms of the reduced form equations. For example, reduced form equations are useful for out-of-sample prediction or forecasting future values of the dependent variable, or for determining which factors are significant without regard to sign or directional effects. If, however, the primary goal is to perform a descriptive analysis of the factors (i.e. independent variables) in the economic process, the researcher can interpret the size and sign of the estimated coefficients using the theoretically determined structural form equations.[3]

The different purposes for the reduced and structural forms for empirical analysis are typically inconsequential since most researchers can specify the structural system of equations. If the simultaneity cannot be completely specified because of an inability to observe a variable or proxy, then estimating and interpreting the sign on coefficients is problematic.[4] In this case, however, the researcher may still be able to predict a future or out-of-sample values of the dependent variable, or determine whether a factor is significant (without regard to sign) by assuming the model is at least a reduced form equivalent.

We use the municipal auditor reputation/fee model from Copley et al. (1995) to demonstrate the development of reduced form equations. Auditor *REPUTATION* and the audit *FEE* are both endogenously determined.[5] Theoretically, the demand and supply of auditor reputation, proxied by auditor size and audit fees, respectively, are determined simultaneously in the market for municipal audit services. Including the auditor's fee as a determinant of the choice of auditor reputation (and vice versa), while theoretically appropriate, gives rise to simultaneity bias in regression estimation unless corrected.

Based on previous research, and a demand and supply theory for auditor reputation and fees, Copley et al. (1995) develop the following structural form equations:

$$\begin{aligned}\text{LN(FEE)} &= \alpha_0 + \alpha_1\text{REPUTATION} + \alpha_2\text{LN(POPUL)} + \alpha_3\text{SERVICES} \\ &\quad + \alpha_4\text{DEBT} + \alpha_5\text{RATING} + \alpha_6\text{SEASON} + e_1 \qquad (1)\end{aligned}$$

$$\begin{aligned}\text{REPUTATION} &= \beta_0 + \beta_1\text{LN(FEE)} + \beta_2\text{TAXSHARE} + \beta_3\text{PROP\_TAX} \\ &\quad + \beta_4\text{NEWDEBT} + \beta_5\text{RATING} + \beta_6\text{MANAGER} + \beta_7\text{LN(POPUL)} \\ &\quad + \beta_8\text{MAYORTURN} + \beta_9\text{B8OFFICE} + e_2 \qquad (2)\end{aligned}$$

The problem of explanatory endogeneity exists because REPUTATION is a right-hand-side variable in the Eq. (1) and a dependent variable in Eq. (2), while LN(FEE) is a right-hand-side variable in Eq. (2) and a dependent variable in Eq. (1).

The first stage of the estimation process is to regress the reduced form equations. Each dependent variable in the system is regressed on all exogenous and predetermined variables in the system. The reduced form equations are as follows:

$$\begin{aligned}\text{LN(FEE)} &= \pi_{10} + \pi_{11}\text{SERVICES} + \pi_{12}\text{LN(POPUL)} + \pi_{13}\text{DEBT} \\ &\quad + \pi_{14}\text{RATING} + \pi_{15}\text{SEASON} + \pi_{16}\text{TAXSHARE} + \pi_{17}\text{PROP\_TAX} \\ &\quad + \pi_{18}\text{NEWDEBT} + \pi_{19}\text{B8OFFICE} + \pi_{110}\text{MANAGER} \\ &\quad + \pi_{111}\text{MAYORTURN} + v_1 \qquad (3)\end{aligned}$$

$$\begin{aligned}\text{REPUTATION} &= \pi_{20} + \pi_{21}\text{SERVICES} + \pi_{22}\text{LN(POPUL)} + \pi_{23}\text{DEBT} \\ &\quad + \pi_{24}\text{RATING} + \pi_{25}\text{SEASON} + \pi_{26}\text{TAXSHARE} + \pi_{27}\text{PROP\_TAX} \\ &\quad + \pi_{28}\text{NEWDEBT} + \pi_{29}\text{B8OFFICE} + \pi_{210}\text{MANAGER} \\ &\quad + \pi_{211}\text{MAYORTURN} + v_2 \qquad (4)\end{aligned}$$

Note that the exogenous explanatory variables are the same in both Eqs (3) and (4). These regression equations produce the predicted values of LN(FEE) and REPUTATION for use as instruments in place of the explanatory endogenous variables in Eqs (1) and (2), respectively.

3.4. Limited Dependent Variables in the System of Equations

A special consideration is whether the system of equations includes qualitative or categorical dependent variables. A significant body of research exists to show

that it is not necessary to have continuous, scale variables to consistently estimate simultaneous equations. Heckman (1978) provides a 2SLS estimation procedure to accommodate the use of qualitative or categorical variables. Amemiya (1978) shows that Heckman's analysis is not efficient and offers an alternative method for incorporating qualitative variables into a 2SGLS model of simultaneous equations. Although Amemiya's procedure is more efficient than that proposed by Heckman (1978), it may overstate the significance of coefficient estimates in small samples. In this case, Deis and Hill (1998) recommend a bootstrap method to obtain unbiased standard errors when Amemiya's (1978) 2SGLS method is used. Overcoming simultaneity bias and addressing estimation problems with limited dependent variables in regressions are important advances since many topics in government accounting include both of these characteristics.

4. SPECIFICATION TESTING FOR SIMULTANEITY BIAS

To illustrate specification testing for simultaneity bias, we test Copley et al.'s data. Since simultaneity is theoretically developed, the use of this data will demonstrate the performance of our test statistics in measuring simultaneity bias. Hausman (1978) develops a test statistic that diagnoses simultaneity bias for linear models with continuous dependent variables. The process of conducting Hausman's test in a traditional simultaneous equations framework (i.e. where all equations are linear) is exactly the same as that described starting with Step 2 below. Step 1, the transformation to a linear model, is obviously not necessary in this case.

Two modifications to Hausman's approach can be made to test systems of equations that include nonlinear models or qualitative dependent variables. The first employs an artificial regression technique suggested by Davidson and MacKinnon (1993) which transforms the nonlinear variables into linear variables, making the usual linear test statistics, such as Hausman's, applicable. The second modification to Hausman's approach relies on evidence provided by Blundell and Smith (1986) that conditional maximum likelihood estimation can be used to perform a Hausman-like test statistic for simultaneity bias. The procedures for these two modifications to Hausman's test are listed below. Using Copley et al.'s data, we demonstrate both tests to: (1) illustrate the consistency between the specification tests; and (2) show that Blundell and Smith's test is easier to perform. The use of an artificial regression adds another step to the specification test (namely, the step in creating the artificial regression) that also has the potential to amplify specification error.

4.1. Hausman Specification Test Using Artificial Regression

Step 1: "Linearize," or create an artificial regression by transforming all variables in the nonlinear model (note that probit, logit, and tobit are nonlinear functions). Transformation functions for binary response models can be found in the econometrics literature (see Davidson & MacKinnon, 1993).[6]

Artificial regressions are simply linear regressions used as calculating devices. In models such as probit, the error term will not be homoscedastic and normally distributed. Yet the error term must exhibit these properties if estimates are to be efficient and inferences valid. The artificial regression provides an asymptotically consistent covariance matrix useful for specification testing. The following details how an artificial regression of Copley et al.'s probit Eq. (2) for auditor reputation is constructed. For ease of exposition, Eq. (2) is restated using matrix algebra (matrices and vectors are shown in bold face) as a nonlinear regression model. Rewriting, the probit equation becomes:

$$\text{REPUTATION} = F(\boldsymbol{\beta}\mathbf{X}) + e,$$

Where $\boldsymbol{\beta}$ is the parameter vector, $\mathbf{X}$ is the matrix of right-hand-side variables [LN(FEE), TAXSHARE, PROP_TAX, NEWDEBT, RATING, MANAGER, LN(POPUL), MAYORTURN, B8OFFICE], and $F(\cdot)$ is the normal cumulative density function. Davidson and MacKinnon (1993) show that a nonlinear regression model of this form yields the following artificial regression:

$$(\text{Var}(\mathbf{bX}))^{-1/2}(\text{REPUTATION} - F(\mathbf{bX})) = (\text{Var}(\mathbf{bX}))^{-1/2} f(\mathbf{bX})\mathbf{bX} + e,$$

where $\mathbf{b}$ is the estimate of the parameter vector $\boldsymbol{\beta}$ from a single equation probit analysis.

While this transformation is tedious, the calculation involves multiplying a scalar times the observed value for each variable. Restating using a superscript T to indicate the variable has been transformed in the artificial regression, the system of structural equations can be rewritten as follows:

$$\begin{aligned}\text{LN(FEE)} &= \alpha_0 + \alpha_1\text{REPUTATION} + \alpha_2\text{LN(POPUL)} + \alpha_3\text{SERVICES} \\ &\quad + \alpha_4\text{DEBT} + \alpha_5\text{RATING} + \alpha_6\text{SEASON} + e_1 \qquad (1)\\ \text{REPUTATION}^T &= g_0 + g_1\text{LN(FEE)}^T + g_2\text{TAXSHARE}^T + g_3\text{PROP\_TAX}^T \\ &\quad + g_4\text{NEWDEBT}^T + g_5\text{RATING}^T + g_6\text{MANAGER}^T + g_7\text{LN(POPUL)}^T \\ &\quad + g_8\text{MAYORTURN}^T + g_9\text{B8OFFICE}^T + e_3 \qquad (2^T)\end{aligned}$$

Step 2. Use OLS to estimate the reduced form Eqs (1) and (2^T) and store the residual values.

Step 3. Use OLS to estimate the structural form and include the residuals from Step 2 above as additional right-hand-side variables (shown as W_1 and W_2 below).

$$\begin{aligned}\text{LN(FEE)} &= h_0 + h_1\text{REPUTATION} + h_2\text{LN(POPUL)} + h_3\text{SERVICES} \\ &\quad + h_4\text{DEBT} + h_5\text{RATING} + h_6\text{SEASON} + h_7 W_1 + e_4 \qquad (1a)\\ \text{REPUTATION}^T &= l_0 + l_1\text{LN(FEE)}^T + l_2\text{TAXSHARE}^T + l_3\text{PROP\_TAX}^T \\ &\quad + l_4\text{NEWDEBT}^T + l_5\text{RATING}^T + l_6\text{MANAGER}^T + l_7\text{LN(POPUL)}^T \\ &\quad + l_8\text{MAYORTURN}^T + l_9\text{B8OFFICE}^T + l_{10} W_2 + e_5 \qquad (2^T a)\end{aligned}$$

Step 4. The null hypothesis of "no simultaneity bias" is tested by the significance of the slope coefficient on the residual variable, W_i. A significant coefficient indicates the presence of simultaneous equations bias. In our tests of Copley et al.'s data, the coefficients on residual variables are significant at traditional levels (t-statistics on the coefficients for W_1 and W_2 are -2.33 and -3.10, respectively), indicating that simultaneity bias exists in these equations.

4.2. Blundell and Smith Specification Test Using Conditional Maximum Likelihood Estimation

Application of Blundell and Smith's (1986) specification test is straightforward since there is no need for an artificial regression.

Step 1. Use OLS on the continuous dependent variable and conditional maximum likelihood on the limited dependent variable to estimate the reduced form equations and store the residual values. The structural form of the model is exactly the same as Eqs (1) and (2) above.

"Conditional" maximum likelihood refers to the nature of the independent variables. If the independent variables in a probit choice analysis are attributes of the "choice" instead of the "chooser," then the analysis is appropriately described as "conditional" probit (Maddala, 1983). For example, audit fees are considered an attribute of the choice of auditor (i.e. Big 5 charge fees that are higher than non-Big 5), whereas the municipal population as a size proxy is considered an attribute of the chooser. If there is at least one independent variable that is an attribute of the choice, then the entire model is considered conditional maximum likelihood.

Step 2. Use OLS or conditional maximum likelihood to estimate the respective structural form and include the residuals from Step 1 above as additional

Table 1. Simultaneous Equations Results Using Blundell and Smith's 2SLS Approach for the Estimation of Reputation and Audit Fees in Municipal Audit Markets.[a]

Predictor Variables[b]	Reduced Form		Structural Form for Simultaneity Specification Test (significance of *W*)		Final Structural Form	
	(1)[c] REPUTATION (Probit)	(2)[c] LN(FEE) (OLS)	(3)[c] REPUTATION (Probit)	(4)[c] LN(FEE) (OLS)	(5)[c] REPUTATION (Probit)	(6)[c] LN(FEE) (OLS)
Intercept	−6.458 (−3.04)***	−3.446 (−4.82)***	−29.364 (−4.17)***	−3.323 (−4.53)***	−28.806 (−4.13)***	−3.395 (−4.78)***
REPUTATION				0.414[d] (1.13)		0.706[e] (2.22)**
LN(FEE)			−6.714[d] (−3.65)***		−6.564[e] (−3.62)***	
SERVICES	−0.229 (−2.64)***	0.027 (0.84)		0.079 (2.33)**		0.079 (2.32)**
LN(POPUL)	0.577 (3.10)***	0.552 (8.68)***	4.257 (3.93)***	0.510 (6.47)***	4.167 (3.90)***	0.511 (6.46)***
DEBT	−0.215 (−1.38)	0.054 (0.97)		0.112 (2.37)**		0.111 (2.36)**
RATING	−0.137 (−1.92)*	0.121 (4.55)***	0.669 (2.96)***	0.153 (5.09)***	0.658 (2.93)***	0.152 (5.05)***
SEASON	−0.563 (−2.31)**	0.068 (0.72)		0.207 (1.83)*		0.206 (1.81)*
TAXSHARE	0.992 (2.59)***	0.177 (1.30)	2.279 (4.01)***		2.238 (3.98)***	
PROP_TAX	1.731 (2.18)**	0.124 (0.44)	2.473 (2.91)***		2.493 (2.94)***	
NEWDEBT	0.407 (1.26)	0.138 (1.19)	1.424 (2.98)***		1.392 (2.94)***	
B8OFFICE	0.059 (0.22)	0.237 (2.33)**	1.668 (3.22)***		1.626 (3.19)***	
MANAGER	0.001 (0.00)	−0.022 (−0.25)	−0.148 (0.62)		−0.136 (−0.58)	
MAYORTURN	0.261 (1.15)	0.043 (0.51)	0.555 (2.22)**		0.547 (2.22)**	
W_i			7.091 (3.81)***	−0.562 (−1.70)*		
N	162	162	162	162	162	162
Rescaled R^2 or Adj R^2	0.296	0.613	0.305	0.619	0.287	0.616
F-stat (ß = 0) or χ^2 (−2LogL)	184.07***	24.15***	182.24***	38.39***	185.478***	43.95***

Comments: Note that the significance of the coefficient on the variable *W* represents the test statistic for the null hypothesis of no simultaneous equations bias. The process of specification testing is as follows. The reduced form estimation (columns (1) and (2) is performed to produce residuals, W_i. W_i is then added to the structural form (columns (3) and (4)) and tested for significance. Columns (5) and (6) are the final estimation results after simultaneity bias has been corrected via 2SLS.

[a] Asymptotic *t*-statistics designated.

[b] All variables, except *W*, defined in Copley et al. (1995).

[c] Values represent coefficient estimates (*t*-statistics).

[d] Actual values of the variables.

[e] Predicted values of the variables from the reduced form estimation.

* For two-tailed *p*-values of 0.10 respectively.

** For two-tailed *p*-values of 0.05 respectively.

*** For two-tailed *p*-values of 0.01 respectively.

right-hand-side variables (shown as W_3 and W_4 below and in Table 1).

$$\text{LN(FEE)} = n_0 + n_1\text{REPUTATION} + n_2\text{LN(POPUL)} + n_3\text{SERVICES} + n_4\text{DEBT} + n_5\text{RATING} + n_6\text{SEASON} + n_7 W_3 + e_6 \quad (1b)$$

$$\text{REPUTATION} = w_0 + w_1\text{LN(FEE)} + w_2\text{TAXSHARE} + w_3\text{PROP\_TAX} + w_4\text{NEWDEBT} + w_5\text{RATING} + w_6\text{MANAGER} + w_7\text{LN(POPUL)} + w_8\text{MAYORTURN} + w_9\text{B8OFFICE} + w_{10} W_4 + e_7 \quad (2b)$$

Step 3. The null hypothesis of "no simultaneity bias" is tested by the significance of the slope coefficient on the residual variable, W_i. A significant coefficient indicates the presence of simultaneous equations bias. The asymptotic *t*-statistics for the coefficients on the residual variables W_3 and W_4 are 3.81 and -1.70 (see Table 1), respectively, suggesting simultaneity bias exists and should be corrected using simultaneous equations estimation.

Note the differences between the procedures in Section A vs. Section B above. In A, an artificial regression is used to ensure that the covariance between the error terms of the equations in the system is zero (a standard assumption in the classical simultaneous equations model (Fomby et al., 1984)). In the Blundell and Smith approach, the use of conditional probit ensures that the covariance between the error terms is zero. Thus, a Hausman-like treatment of the residuals from the reduced form to test for contemporaneous correlation between regressors and the error term is directly possible with conditional probit under the Blundell and Smith approach.

In summary, specification testing using the Hausman test on artificial regressions, or the Blundell and Smith test with conditional maximum likelihood, results in the same inference for Copley et al.'s data: auditor reputation and audit fees are jointly and endogenously determined. Our specification test results are consistent with the argument provided in Copley et al. (1995) that auditor choice and audit fees should be examined within a system of simultaneous equations. Although the primary purpose of using Copley et al.'s data is to show how to perform the Blundell and Smith test, it is worthwhile noting that Copley et al. did not perform a specification test for simultaneous equations bias. Thus, our specification test empirically confirms the analytical framework developed in their paper.[7]

5. SUGGESTIONS FOR FUTURE RESEARCH

Though a number of government accounting researchers have begun to focus on the issue of endogeneity, several possible questions remain. Following is a sample

of the types of issues that may be affected by endogeneity and benefit from the use of simultaneous equations.

Baber (1994) discusses the influence of political competition on governmental reporting. Political competition can be a factor in governmental accounting (disclosure) choices because an elected official's incentives to respond to interest group demands depends on whether there are credible challenges in future elections. However, political competition itself could also be a function of governmental accounting disclosure. That is, accounting disclosures by elected officials might change the political competitive environment or discourage political challengers.

Ingram, Raman and Wilson (1987) summarize the research related to municipal bond ratings. An extension of that research could apply the work of Zeibart and Reiter (1992) to the municipal bond markets. Net interest cost is typically modeled as a function of governmental accounting and auditing measures and bond ratings. However, it is also likely that governmental accounting information has an indirect effect on net interest costs through its influence on municipal bond ratings.

Langsam and Kreuze (1990) found a positive correlation between municipal bond ratings and audit opinions, but did not address causality (Reck, Wilson, Gotlob & Lawrence, 2004). Perhaps there is dual causality, with the audit opinion being based on, and determinant of, the assigned bond ratings.

Studies by Benson, Marks and Raman (1991) and Fairchild and Koch (1998) find that voluntary disclosures by governmental entities decrease net interest costs on municipal bonds. However, it is possible that a government's decision to provide private, voluntary information is motivated by net interest costs.

Other possible examples of variables that are concurrently determined include auditor choice and accounting choices. A decision to use a Big 4 audit firm may be driven by a plan for early implementation of new accounting standards (such as GASB 34). Likewise, the ability to adopt a new standard early may be dependent on audit quality, such as that available within a Big 4 or other specialist audit firm.

6. CONCLUDING REMARKS

In order to draw reasonable conclusions from statistical analyses, it is necessary to assure that none of the assumptions of classical methodologies are violated. Researchers are trained to avoid a number of statistical pitfalls that affect the interpretation of research results, including multicollinearity of explanatory variables, non-normality of data and heterogeneity of the variance in error terms. One additional issue about which researchers need to remain vigilant is the problem of explanatory endogeneity. Often, the questions addressed in

accounting research involve cause and effect that is not unidirectional. The issues in government accounting are no exception and may exhibit similar simultaneity.

In addition to reviewing previous research that exemplifies estimation and control of explanatory endogeneity, we provide examples of specification testing for simultaneity bias in governmental auditing research and demonstrate how to perform two separate specification tests. Simultaneity bias in governmental audit markets is well established in previous studies and confirmed in our tests.

We also suggest some additional research issues in government accounting that may be appropriately addressed by examining a system of simultaneous equations. If simultaneity exists, but is ignored in a research model, the estimation may result in coefficients that are biased and inconsistent, and corresponding inferences that are invalid.

NOTES

1. Other examples include, Ettredge, Simon, Smith and Stone (1994) and Hogan (1997). Ettredge et al. estimate the relation of audit fees (a continuous variable) and the engagement of quarterly reviews (a binary variable). They find that while the performance of timely quarterly reviews reduces audit fees, the audit fee simultaneously influences whether the auditor would undertake quarterly reviews. Hogan finds that in the market for initial public offerings (IPOs), many of the same factors that determine whether to retain a Big 6 audit firm also determine audit cost. Audit cost, in turn, is directly affected by the selection of a Big 6 auditor.

2. The Copley et al. (1994) model uses both continuous (audit fees) and binary (audit quality) variables.

3. Empirical-archival research can be both descriptive and predictive. In this paper, descriptive refers primarily to a focus on the coefficients of the independent variables while predictive refers to a focus on the estimated values of the dependent variable.

4. Since the coefficient on a reduced form independent variable is an algebraic manipulation of the structural form coefficients, the sign cannot be interpreted unless the structural form is known.

5. LN(FEE) is the natural logarithm of the audit fee; REPUTATION is a dichotomous variable coded as 1 if the auditor is a Big 8 firm and zero otherwise; LN(POPUL) is the natural logarithm of the municipality's population; SERVICES is an index of municipal services; DEBT is total per capita long-term debt; RATING is a variable that indicates how far below AAA the general obligation debt's rating is; SEASON indicates whether or not the fiscal year end coincides with audit busy season; TAXSHARE is the total per capita revenue excluding transfers from other governmental units; PROP_TAX is the percentage of local taxes generated by property taxes; NEWDEBT is the total per capita debt issued in the year subsequent to the audit; MANAGER is a variable that indicates whether the municipality has a city-manager government structure; MAYORTURN indicates whether the mayor changed in the next election following the audit; and B8OFFICE indicates whether the audit firm has a local office in the municipality.

6. Transformations for artificial regressions come in many forms. In the case of probit, the variance and the normal density of $\mathbf{X\beta}$ is originally obtained from a probit regression for reuse in the artificial regression estimation: $(\text{Var}(\mathbf{bX}))^{-1/2}(\text{REPUTATION} - F(\mathbf{bX})) = (\text{Var}(\mathbf{bX}))^{-1/2} f(\mathbf{bX})\mathbf{bX} + e$.

7. In columns (5) and (6) of Table 1 we provide the final structural estimation of Copley et al.'s reputation-fee model. Although the sign and significance of our results are qualitatively the same as Copley et al.'s results, minor differences in the estimates are due to the fact that we use Heckman's 2SLS whereas Copley et al. use Amemiya's 2SGLS.

ACKNOWLEDGMENTS

We thank participants at the 2003 American Accounting Association Annual Meeting, the editor, and two anonymous referees for helpful comments. We also appreciate the research support from the School of Management at George Mason University and the GMU Accounting Excellence Fund.

REFERENCES

Amemiya, T. (1978). The estimation of a simultaneous equation generalized probit model. *Econometrica*, *46*(5), 1193–1205.

Baber, W. R. (1994). The influence of political competition on governmental reporting and auditing. *Research in Governmental and Nonprofit Accounting*, *8*, 109–127.

Benson, E. D., Marks, B. R., & Raman, K. K. (1991). The effect of voluntary GAAP compliance and financial disclosure on governmental borrowing costs. *Journal of Accounting, Auditing, and Finance*, *6*(3), 303–319.

Berry, W. D. (1984). *Nonrecursive causal models*. Beverly Hills, CA: Sage.

Blundell, R. W., & Smith, R. J. (1986). An exogeneity test for simultaneous equation tobit models with an application to labour supply. *Econometrica*, *54*, 679–685.

Copley, P. A., Doucet, M. S., & Gaver, K. M. (1994, January). A simultaneous equations analysis of quality control review. *Accounting Review*, *69*(1), 244–256.

Copley, P. A., & Douthett, E. B., Jr. (2002, Spring). The association between auditor choice, ownership retained, and earnings disclosure by firms making initial public offerings. *Contemporary Accounting Research*, *19*(1), 49–75.

Copley, P. A., Gaver, J. J., & Gaver, K. M. (1995, Spring). Simultaneous estimation of the supply and demand of differentiated audits: Evidence from the municipal audit market. *Journal of Accounting Research*, *33*(1), 137–155.

Davidson, R. D., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. New York, NY: Oxford University Press.

Deis, D. R., & Hill, R. C. (1998, Spring). An application of the bootstrap method to the simultaneous equations model of the demand and supply of audit services. *Contemporary Accounting Research*, *15*(1), 83–99.

Demski, J. S., Fellingham, J. C., Ijiri, Y., & Sunder, S. (2002, June). Some thoughts on the intellectual foundations of accounting. *Accounting Horizons*, *16*(2), 157–168.

Elder, R. J., Kattelus, S. C., & Douthett, E. B. (1999). Audit fees and nonaudit fees in the governmental sector: A self-selection analysis. *Research in Government and Nonprofit Accounting, 10*, 65–86.

Ettredge, M., Simon, D., Smith, D., & Stone, M. (1994, September). Why do companies purchase timely quarterly reviews? *Journal of Accounting and Economics, 18*(2), 131–155.

Fairchild, L. M., & Koch, T. W. (1998). The impact of state disclosure requirements on municipal yields. *National Tax Journal, 51*(4), 733–753.

Fomby, T. B., Hill, R. C., & Johnson, S. R. (1984). *Advanced econometric methods*. New York, NY: Springer-verlag.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica, 46*(6), 1251–1271.

Heckman, J. (1978). Dummy endogenous variables in simultaneous equation systems. *Econometrica, 46*(6), 931–960.

Hogan, C. (1997, January). Costs and benefits of audit quality in the IPO market: A self-selection analysis. *Accounting Review, 72*(1), 67–86.

Ingram, R. W., Raman, K. K., & Wilson, E. R. (1987). Governmental capital markets accounting research in account: A review. *Research in Governmental and Nonprofit Accounting, 3*, 111–132.

Ireland, J. C., & Lennox, C. S. (2002, Winter). The large audit firm fee premium: A case of selectivity bias? *Journal of Accounting, Auditing and Finance, 17*(1), 73–91.

Langsam, S. A., & Kreuze, J. G. (1990). The relationship of type of audit opinion to governmental bond ratings. *Arkansas Business & Economic Review, 23*(2), 1–10.

Leftwich, R. W., Watts, R. L., & Zimmerman, J. L. (1981). Voluntary corporate disclosure: The case of interim reporting. *Journal of Accounting Research, 19*, 50–88.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. New York: Cambridge University Press.

Reck, J. L., Wilson, E. R., Gotlob, D., & Lawrence, C. M. (2004). Governmental capital markets research in accounting: A review, extension, and directions for future research. *Research in Government and Nonprofit Account, 11*, 1–33.

Yatchew, A., & Griliches, Z. (1984). Specification error in probit models. *Review of Economics and Statistics, 66*, 134–139.

Zeibart, D. A., & Reiter, S. A. (1992, Fall). Bond ratings, bond yields and financial information. *Contemporary Accounting Research, 9*(1), 252–282.

AN EMPIRICAL INVESTIGATION OF MISAPPROPRIATION RED FLAGS IN GOVERNMENTS: SAS 82 AND BEYOND

Carolyn Strand Norman, Sandra T. Welch,
Florence C. Sharp and Sarah A. Holmes

ABSTRACT

This paper reports the results of a survey mailed to members of the Association of Government Accountants that identifies red flags commonly associated with frauds perpetrated against governmental entities and contrasts these to the red flags enumerated in SAS 82. The frauds detailed by the respondents were perpetrated against all levels of government and spanned the period 1990 to 1998. Not unexpectedly, computers were employed in many schemes (47 of 178 cases). Fraudulent disbursements were utilized in 80% of the federal cases reported, while the theft of cash and other assets were more common at the state (44%) and local (52.6%) levels. Dollar losses averaged $1.3 million at the federal level, $687 thousand at the state level, and $259 thousand at the local level.

The results of the study provide support for applying the risk assessment requirements of SAS 82 in the governmental environment. Specifically, weak or non-existent internal controls, such as lack of management oversight (109 of 178 cases) and inadequate segregation of duties (104 of 178 cases) often create an environment that encourages individual fraudulent behavior. Red flags related to personnel conditions, accounting record discrepancies,

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 151–182

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11008-2

and conflicting/missing evidential material also serve as precursors of fraud. Important red flags specifically cited in SAS 82 included inadequate or incomplete accounting for transactions (68 of 178 cases), missing documentation (65 of 178 cases), observed lifestyle changes (51 of 178 cases), and personnel with known financial problems (46 of 178 cases). Certain red flags not identified or even implied in SAS 82 may also serve as warning signals that the misappropriation of governmental assets may be occurring. These include altered documentation (65 of 178 cases), delays in reconciliations (53 of 178 cases), and defensive behavior by the perpetrator when confronted with questions (37 of 178 cases).

Only one-fourth of the government frauds reported in this study were discovered through the audit function. Suspicions expressed by employees led to the discovery of many fraudulent acts (77 of 178 cases). Knowledge of our findings should enable concerned parties within government to create an organizational atmosphere conducive to the detection of fraud in a more timely manner, and the creation of an internal control system that would offer greater fraud deterrence.

1. INTRODUCTION

Fraud has been extremely costly to governmental organizations. It is not unusual to read news about heath care providers submitting fraudulent Medicare claims, about unqualified individuals collecting welfare payments, or about public employees embezzling public money. A recent General Accounting Office (GAO) report estimated that as much as one-tenth of health expenditures nationwide are lost to fraud and abuse, which amounts to approximately $100 billion annually (GAO, 1997; Thompson, 1992). In addition to monetary losses, fraud has public policy implications. It diverts budget resources from providing public services, undermines governments' attempts at reform, and erodes public confidence in government (Welch et al., 1997, p. 38).

From a public policy standpoint, the high cost of fraud provides a strong argument for auditors to devote resources to its detection. Although auditors traditionally have been associated with fraud detection, for decades the American Institute of Certified Public Accountants (AICPA) avoided accepting direct responsibility for detection of fraud through the Generally Accepted Auditing Standards. It was not until 1997 that Statement on Auditing Standards No. 82 (SAS 82) *Consideration of Fraud in a Financial Statement Audit* made a clear statement that planning for detection of fraud should be covered in an audit of an organization's financial statements (AICPA, 1997). SAS 82 identifies several red

flags that may indicate fraud, but there is little evidence to suggest whether the list is particularly suited to detection of fraud in governmental entities.

The purpose of this study is to identify red flags associated with frauds taking place in governmental entities, and to compare these risk factors to the red flags currently enumerated in SAS 82. A survey of governmental accountants was used to obtain data regarding types of frauds that have occurred in governmental entities. For each fraudulent activity described, the survey gathered information on the detection of the fraud, the internal control structure existing in the victim entity, and any red flags known to be present at the time of the defalcation.

We begin with a brief history of the AICPA's position on the auditor's responsibility for the detection of fraud, followed by a description of the key provisions of SAS 82. Characteristics of governmental organizations that may be relevant to the potential for fraud in that sector of the economy are then discussed. This is followed by an identification of the research questions guiding this study and a description of the methodology used. We then present and discuss the research results. Finally, the conclusions describe the potential contribution of the research.

2. BACKGROUND

2.1. Auditors' Responsibility for Fraud Detection

One might argue that the AICPA's first attempt to deal with fraud occurred in 1939 with its issuance of Statement on Auditing Procedures No. 1: *Extensions of Auditing Procedures* (SAP 1). Specifying the need to confirm receivables and physically inspect inventories, SAP 1 was issued in direct response to the discovery of the McKesson and Robbins fraud in 1938 and the resulting negative publicity against both the auditing firm in particular, and the auditing profession as a whole.

Several subsequent standards (SAS 16: *The Independent Auditor's Responsibility for the Detection of Errors and Irregularities* (1977); SAS 53: *The Auditor's Responsibility to Detect and Report Errors and Irregularities* (1988); SAS 54: *Illegal Acts by Clients* (1988); SAS 82: *Consideration of Fraud in a Financial Statement Audit* (1997)) have further addressed the auditor's responsibility for detection of fraud. Elements of SAS 16, SAS 53, and SAS 82 are summarized in Table 1.[1]

In the aftermath of several widely publicized criminal prosecutions,[2] the AICPA issued SAS 16 to provide guidance to auditors concerning their search for material errors and irregularities. Although this Standard did not clearly establish responsibility for fraud detection, it did identify three risk factors: failure to correct material weaknesses in internal controls, high turnover in key positions, and understaffing in accounting and financial areas.

Table 1. Comparison of SAS 16, 53, and 82.

	SAS 16	SAS 53	SAS 82
Date released	1977	1988	1997
Scope	Deals with errors and irregularities in a financial statement audit.	Same as SAS 16	Deals with auditor's consideration of fraud in a financial statement audit.
Definitions	Errors and irregularities. Irregularities include both fraudulent financial reporting and misappropriations of assets (para. 2, 3).	Same as SAS 16	Expanded description of fraud; covers both fraudulent financial reporting and misappropriations (para. 3–10).
Risk assessment responsibility		To assess the risk that errors and irregularities may cause the financial statements to contain a misstatement.	Specifically assess the risk of material misstatement of the financial statements due to fraud.
Detection responsibility	To plan the audit examination to search for errors or irregularities that would have a material effect on the financial statements.	Based on assessment, design the audit to provide reasonable assurance of detecting errors and irregularities that are material to the financial statements.	Based on assessment, design audit procedures to be performed. Should consider fraud risk factors that relate to fraudulent financial reporting and misappropriation of assets (para. 16, 18). Clears up some ambiguities; uses "fraud" rather than "irregularities"; Must plan and perform the audit to obtain reasonable assurance about whether the financial statements are free of material misstatement, whether caused by error or fraud.

Risk factors	Company does not correct material weaknesses in internal control; Key positions have high turnover; Accounting and financial functions understaffed.	Identifies 16 factors/characteristics in these categories: – Management – Operating and industry – Engagement Also identifies nine risk factors that may influence the assessment of audit risk at the balance or class level.	Many more examples of risk factors than contained in SAS 53. For misappropriation of assets: Susceptibility of assets to misappropriation, and controls.
Other conditions	None	None	13 examples of other conditions that may be identified during field work that change or support the risk assessment: Discrepancies in accounting records; Conflicting or missing evidential matter; Problematic relations between auditor and client.
Risk response	No examples	No examples	Examples of responses to: Risk at the account balance, class of transactions, and assertion level; Risk of misappropriation of assets.
Professional skepticism	Should plan and perform audit examination with an attitude of professional skepticism.	Discusses the concept and describes professional skepticism in planning and performing the audit.	An expanded discussion included in AU Section 230, *Due Professional Care in the Performance of Work*, to emphasize professional skepticism throughout the audit.

Table 1. (*Continued*)

	SAS 16	SAS 53	SAS 82
Response when error, irregularity, or fraud is suspected	If auditor believes material errors or irregularities exist: Discuss with appropriate level of management, at least one level above those involved. Should attempt to obtain sufficient evidential matter to determine whether errors or irregularities do exist.	Same as SAS 16	Immaterial effect: Refer to an appropriate level of management at least one level above those involved; Be satisfied that implications for other aspects of the audit have been adequately considered. Material effect: Consider implications for other aspects of the audit; Discuss with appropriate level of management, at least one level above those involved; Attempt to determine whether fraud exists and, if so, its effect.
Documentation required	None	None	Risk factors identified as present and the auditor's response to those risk factors; If other risk factors are identified during the audit that cause the auditor to believe that an additional response is required, the auditor should document those risk factors and any further response the auditor deems appropriate.

Inquiries		Auditor should inquire of management (1) to obtain their understanding regarding the risk of fraud in the entity and (2) to determine whether they have knowledge of fraud that has been perpetrated on or within the entity (para. 13). If the entity has established a program to prevent, deter, and detect fraud, the auditor should inquire of those persons overseeing such programs as to whether the program has identified any fraud risk factors (para. 24).
Communications	If after discussions with management, the auditor continues to believe material error or irregularity exists, determine if Board of Directors or Audit Committee is aware of circumstances. Irregularities involving senior management should be reported directly to audit committee.	When the auditor has identified risk factors that have continuing control implications, the auditor should consider whether these risk factors represent reportable conditions that should be communicated to senior management and the audit committee.

When massive savings and loan failures, labeled the costliest financial disaster in United States history (Epstein, 1993), occurred in the 1980s, the news media and the public questioned whether auditor negligence had precipitated the crisis. In response to these concerns, the AICPA, the American Accounting Association, the Financial Executives Institute, the Institute of Internal Auditors, and the National Association of Accountants formed and funded the National Commission on Fraudulent Financial Reporting (Treadway Commission). From 1985 to 1987 the Treadway Commission studied all aspects of fraudulent financial reporting, including the role of auditors. These deliberations culminated in the issuance of two auditing standards: SAS 53 and SAS 54. SAS 53 required the auditor, first, to assess the risk that errors and irregularities might cause material misstatement and, then, to design the audit to provide reasonable assurance that material errors and irregularities would be detected. SAS 53 identified sixteen management, engagement, operating, and industry characteristics that increase the risk that material errors and irregularities might occur and nine factors that might affect the audit risk assessment at the balance or class level. SAS 54 provided guidance for the auditor when violations of laws or regulations are encountered in the course of an audit.

Some groups expressed concern that SAS 53 did not go far enough in defining auditors' responsibility for the detection of fraud. In March 1993, the AICPA's Public Oversight Board made a number of recommendations regarding fraud and called for greater professional skepticism (Mancino, 1997). Later that year, the AICPA formed a task force to look at SAS 53 more critically. The Auditing Standards Board finally decided that a standard focusing solely on fraud, including both financial statement fraud and misappropriation of assets, should be developed. While these deliberations were taking place, the Committee of Sponsoring Organizations of the Treadway Commission (COSO) sponsored another study of fraudulent financial reporting in publicly listed companies. This study (Beasley et al., 1997) identified higher risks of fraud by upper management and in small companies. Study results highlighted the role of the control environment in countering fraud.

Finally, SAS 82 was issued in February 1997. SAS 82 clarified Section 110 of the auditing standards by replacing the word "irregularity" with "fraud":

> The auditor has a responsibility to plan and perform the audit to obtain reasonable assurance about whether the financial statements are free of material misstatement, whether caused by error or fraud (p. 29).

The Standard provides guidance to auditors in carrying out this responsibility by identifying and classifying risk factors (red flags) relating to fraudulent financial reporting and to misappropriation of assets. The red flags related to the misappropriation of assets, the concern of this study, are linked to two categories

of circumstances: the susceptibility of assets to misappropriation and internal control problems. The Standard lists four red flags regarding the susceptibility of assets to misappropriation. These risk factors "pertain to the nature of an entity's assets and the degree to which they are subject to theft" (SAS 82, para. 18a). Eight red flags are provided pertaining to the internal control arena. These risk factors "involve the lack of controls designed to prevent or detect misappropriations of assets" (SAS 82, para. 18b). The Standard elaborates on these 12 red flags by enumerating additional factors that should trigger the concern of the auditor (SAS 82, para. 20–25). The red flags identified in SAS 82 are summarized in Table 2.

2.2. Special Industry Issues

Albrecht et al. (1995, p. 55) note that "... with fraud, it is not initially certain that a crime has been committed. Only fraud symptoms, red flags, or indicators are seen." Red flags can be used to create fraud risk checklists for the auditor. However, research by Pincus (1989) on audit decision aids illustrates the importance of including the appropriate red flags on checklists designed to facilitate auditors' decisions. The red flags listed in SAS 82 are a combination of risk factors from previous standards and audit risk alerts derived from auditors' specific experiences. The red flags enumerated may not be equally applicable to all audits and thus may not necessarily apply to the public sector. We investigate the adequacy of the SAS 82 red flags in helping auditors evaluate the potential for fraud in a governmental environment.

It would also be useful to ascertain whether other red flags not listed in SAS 82 are relevant to the public sector. Albrecht et al. (1995, p. 96) point out that "accounting, control, and organizational anomalies often provide symptoms that fraud is being perpetrated." Loebbecke et al.'s (1989) study summarizing auditors' experience with irregularities reported that 42.7% of all irregularities reported by Peat Marwick (KPMG) partners occurred in revenues, accounts receivable, and inventory. Only 16.4% of the frauds involved cash. The largest industry representations in the Loebbecke et al. study were banking (19.0%), manufacturing (14.1%), merchandising (12.5%), and savings and loans (12.2%). In contrast, research by Green and Choi (1997) on the financial services sector pointed out that inventories are largely irrelevant in that sector of the economy. Instead, they observed that the three most frequently misstated accounts were loan loss reserves, investments, and loan receivables, respectively. These findings suggest that the factors associated with the risk of fraud and the red flags that might alert one to the presence of fraud are not the same in all industries. There may be predominant types of fraud and/or

Table 2. Red Flags Identified in SAS 82.

Risk factors related to susceptibility of assets to misappropriation
- Large amount of cash on hand or processed
- Inventory characteristics, such as small size, high value, or high demand
- Easily convertible assets, such as bearer bonds, diamonds, or computer chips
- Fixed asset characteristics, such as small size, marketability, or lack of ownership identification

Risk factors related to controls
- Lack of management oversight (for example, inadequate supervision or monitoring of remote locations)
- Lack of job applicant screening procedures relating to employees with access to assets susceptible to misappropriation
- Inadequate record keeping with respect to assets susceptible to misappropriation
- Lack of segregation of duties or independent checks
- Lack of appropriate system of authorization and approval of transactions (for example, in purchasing)
- Poor physical safeguards over cash, investments, inventory, or fixed assets
- Lack of timely and appropriate documentation for transactions (for example, credits for merchandise returns)
- Lack of mandatory vacations for employees performing key control functions

Personnel Conditions Related to Misappropriation of Assets
- Anticipated future employee layoffs that are known to the workforce
- Employees with access to susceptible assets who are known to be dissatisfied
- Known unusual changes in behavior or lifestyle of employees with access to assets susceptible to misappropriation
- Known personal financial pressures affecting employees with access to assets susceptible to misappropriation

Other considerations related to risk factors in assessing the risk of material misstatement due to fraud

General
- Risk factors individually or in combination
- Size, complexity, and ownership characteristics of entity
- Mitigating controls in place
- Effectiveness of controls

Discrepancies in the accounting records
- Transaction not recorded in complete or timely manner or improperly recorded as to amount, accounting period, classification, or entity policy
- Unsupported or unauthorized balances or transactions
- Last-minute adjustments by the entity that significantly affect financial results

Conflicting or missing evidential matter
- Missing documents
- Unavailability of other than photocopied documents when documents in original form are expected to exist
- Significant unexplained items on reconciliations
- Inconsistent, vague, or implausible responses from management or employees arising from inquiries or analytical procedures

Table 2. (*Continued*)

Unusual discrepancies between the entity's records and confirmation replies
Missing inventory or physical assets of significant magnitude
Problematic or unusual relationships between auditor and client
Denied access to records, facilities, certain employees, customers, vendors, or others from whom audit evidence might be sought
Undue time pressures imposed by management to resolve complex or contentious issues
Unusual delays by the entity in providing requested information
Tips or complaints to the auditor about fraud

red flags not identified in SAS 82 that need to be highlighted in the public sector. Identification of such elements would allow auditors to design audits to recognize such activities and would improve their subsequent fraud risk assessments.

2.3. Characteristics of Governmental Entities

2.3.1. Organizational Factors

There are two aspects of government that set it apart from for-profit entities. First, although both businesses and governments provide services, government recipients typically do not pay directly for the public goods received. Taxpayers, grantors, and other governmental units usually provide resources to the service provider, with amounts often established by legislative fiat. Since service recipients do not enter into exchange transactions with the service provider, they are less likely to provide the level of control typically exercised by customers in the private sector. In such cases, red flags related to competition and market saturation are largely irrelevant.

Second, the concept of accountability is operationalized differently in the private sector than it is in government. In the private sector, management generally has the authority to choose how to deploy its resources. In contrast, most governmental financial activity is budget-driven, and the budget, once approved, is a legally binding document. Thus, red flags related to budgetary or taxpayer concerns are highly relevant in the public sector. Specifically, the misappropriation of assets in a governmental context is expanded to include the disbursement of resources for non-appropriated purposes. For example, many governmental assets are restricted for particular activities or purposes, and the legislative body has no authority to budget their expenditure for any other purpose. The federal government may award a city a grant to be used for hiring police officers, or voters may approve a tax levy for capital improvements. No matter how worthwhile another project might be, designated funds can be used only for the specified purpose.

2.3.2. Public Policy Issues

Certain recently implemented public policy decisions, such as downsizing and devolution, have increased the potential for fraud in governmental organizations. Though not unique to the public sector, downsizing, which is accomplished through reductions in personnel and/or resources, has been identified by federal financial managers as a risk factor (Steininger et al., 1997). Downsizing can eliminate levels of internal control that previously existed and can diminish the entity's ability to appropriately manage its activities.

There has also been political pressure to initiate devolution, which delegates entire functions to a lower governmental level. The decision to transfer responsibility from the federal to the state level, or from the state to the local level, is sometimes undertaken to move the entity providing the service closer to the constituents who use or desire the service, thus creating greater public responsiveness to emerging problems. However, lack of experience and cost constraints may prevent the lower governmental level from implementing the appropriate internal controls, thus increasing the risk of misappropriation of assets.

The fact that the government environment is so different from the corporate environment suggests that the risk factors identified in SAS 82, while useful, may require modification for the public sector. Government auditors may be required to consider a more complex or comprehensive set of risk factors if they are to do a complete assessment of the possibilities for fraud and the risk of fraud. Certainly, some risk factors may be more important in government than in business settings. The purpose of this study is to provide some evidence supporting the relevance for the governmental environment of some of the risk factors contained in SAS 82 and to expand the list of risk factors to include others of special concern to government.

3. RESEARCH QUESTIONS

The following three research questions are addressed in this study. SAS 82 offers a set of red flags applicable to misappropriation of assets. The first question addresses whether the red flags identified in SAS 82 are useful to government entities.

Research Question 1. Which of the red flags in SAS 82 facilitate the identification of misappropriation frauds in governments?

As already discussed, the governmental sector of the economy has characteristics that differ from the private sector with respect to both revenue generation and expenditures. Misappropriation of assets in the government environment, therefore, may display different characteristics. The second research question addresses this issue.

Research Question 2. Are any red flags not included in SAS 82 relevant to misappropriation of assets in the governmental arena?

Different levels of government may possess differing strengths and weaknesses in protecting their assets from misappropriation. To the extent that lower levels of government are unprepared for increased levels of responsibility, circumstances may be more conducive to fraud at lower levels. The third research question addresses this concern.

Research Question 3. Do red flags related to asset misappropriation differ at the federal, state, and local levels of government?

4. METHODOLOGY

In order to gather data to address the issues discussed above, a five-page instrument was designed by the researchers and mailed to all national members of the Association of Government Accountants. The instrument asks each respondent to describe, in detail, an instance of fraud with which (s)he is personally familiar. The description includes characteristics of the fraud victim and the perpetrator(s); the type, size, and duration of the fraud; and how the fraud was discovered. The instrument also asks respondents to identify specific red flags that led to the discovery of the fraud, including both internal control problems and factors indicating susceptibility to misappropriation of assets.

A total of 178 usable responses were returned to the researchers, who in turn coded and recorded the responses for further analysis. Table 3 summarizes the instruments returned.

Table 3. Demographic Information from a Sample of 178 Members of the Association of Government Accountants Completing a 1998 Survey on Public Sector Fraud.

	Total	Level of Government		
		Federal	State	Local
Cases reported				
Number of fraud cases reported	178	48	52	78
% of reported cases		27.0	29.2	43.8
Demographics of population sampled				
AGA members employed in federal, state, or local government	12,369	5,468	4,741	2,160
% of members		44.2	38.3	17.5

While we cannot argue that we achieved a normal response rate, we assert that the relative rarity of the event being examined (misappropriation of assets), coupled with the detailed knowledge required for each case, explain, in part, our low returns.[3] Most likely, respondents only reported on frauds that they had actually investigated. Indeed, 110 individuals cited the recency, completeness, or their familiarity with case particulars as the reason this fraud was selected for response. Irrespective of the response rate, details of 178 instances of governmental fraud paint a picture not previously available. We believe that this wealth of information fills a gap in the literature. The responses are summarized, both for overall information and by type of government in which the fraud occurred (federal, state, or local). Differences among nominal variables across the three types of governments are analyzed through the construction of contingency tables and calculation of the Chi-squared statistic. Analysis of variance (ANOVA) is used to test differences in means.

5. RESULTS AND DISCUSSION

The respondents provided a considerable amount of information about each reported fraud. Details pertaining to both the victims and the perpetrators of the frauds, the schemes, the methods of discovery, and the red flags that created concern are discussed in the following five sections.

5.1. The Victims of the Frauds

The characteristics of the victim organizations are summarized in Table 4.

Of the 178 frauds reported in this study, 48 occurred in the federal government, 52 in state government, and 78 in the local environment. Most are relatively current: 138 occurred between 1990 and 1998. Although 39 states and the District of Columbia are represented in the sample, over half the reported frauds (92) occurred in California, Ohio, Washington, Michigan, Georgia, New York, Virginia and Washington D.C. Relatively few of the entities at any level had audit committees when the fraud occurred (federal = 18.2%; state = 23.5%; local = 19.2%). However, a significant number of the defrauded federal entities (77.3%) reported the presence of an internal audit function. In contrast, only 45.1% of the state organizations and 35.9% of the local entities had internal auditors. This suggests that, at the federal level, resolution of and follow-up on internal audit findings may not be accomplished in a timely manner. Alternatively, disruptions created by changes in either the services provided or the personnel providing the services may have created an environment conducive to fraud, particularly at the federal level.

Table 4. Characteristics of 178 Victims of Fraud in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

	Total (178)	Level of Government		
		Federal (48)	State (52)	Local (78)
Mean budget size ($ millions)***	200.4	434.6	204.1	110.6
Victims had				
Audit committee	$n = 35$	18.2%	23.5%	19.2%
Internal audit function***	85	77.3	45.1	35.9
Victims experienced				
Personnel changes**				
Decreased	31	28.6	22.4	10.5
Stayed the same	112	47.6	65.3	78.9
Increased	24	23.8	12.2	10.5
Did not respond to survey item	11			
Changes in services provided**				
Decreased	8	11.9	4.1	1.3
Stayed the same	121	54.8	79.6	77.6
Increased	38	33.3	16.3	21.1
Did not respond to survey item	11			

** Differences between levels of government are significant at $p = 0.05$.
*** differences between levels of government are significant at $p = 0.01$.

5.2. The Perpetrators of the Frauds

Table 5 summarizes the characteristics of the perpetrators of the reported frauds.

Insiders acting alone committed almost two-thirds of the frauds. Employees perpetrated 40.4% of the misappropriations at the state level, while managers were more active at the local level (42.3%). Possibly reflecting the greater presence of internal auditors (and accompanying controls), collusion occurred more frequently at the federal level (42.5%, versus 23.0% for state frauds and 24.4% for local frauds). With a reported average tenure exceeding 8 years, perpetrators were probably knowledgeable of any opportunities provided in their work setting by control weaknesses.

5.3. Description of the Fraud Schemes

All 178 frauds reported were cases of asset misappropriation. Characteristics of these schemes are summarized in Table 6.

Table 5. Characteristics of the Perpetrators of 178 Incidents of Fraud in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Type of perpetrator				
Employee***	$n = 55$	19.1	40.4	32.1
Manager***	57	23.4	25.0	42.3
Vendor/service recipient***	14	14.9	11.5	1.3
Insider/insider collusion***	29	25.5	11.5	14.1
Insider/outsider collusion***	22	17.0	11.5	10.3
Tenure of internal perpetrators				
Years in position (mean in years)	8.4	9.0	7.5	8.6

*** Differences between levels of government are significant at $p = 0.01$.

As might be expected, the magnitude of the losses varied significantly across governmental levels, with federal organizations losing an average of $1,300,760, state organizations losing $687,013, and local entities losing $259,381. The majority (105) of all schemes involved fraudulent disbursements. Perpetrators stole cash receipts in 62 of the reported cases and other assets in 11 more instances. Misappropriations in the federal sector were more likely to occur in the disbursements cycle (81.3%). In contrast, state and local fraudulent activity was more evenly distributed between receipts/assets (state = 44.2%; local = 52.6%) and expenditures (state = 55.8%; local = 47.4%).

Significantly more federal-level fraud perpetrators used the computer in their schemes (37.5%) than did state-level (19.2%) or local-level (24.4%) perpetrators. Over three-fourths of all computer frauds involved the modification of input data. This ploy was particularly evident at the federal level (83.3%), versus state (70.0%) and local (73.7%). Only 24.7% of the respondents said the type of fraud they reported occurred "only rarely" or was a "one-time event," while 36.5% thought it happened "occasionally," and another 32.6% judged the occurrence to be "somewhat frequent" to "frequent."

The average fraud scheme operated undetected for almost three years, with no significant differences noted across levels of government. The most frequent explanation given for how the schemes were successfully concealed was that no oversight was provided for the activities (federal = 40.9%; state = 50.0%; local = 40.0%). The preparation of false documentation (federal = 40.9%; state = 26.0%;

Table 6. Characteristics of 178 Fraud Schemes in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Type of fraud				
Fraudulent disbursements**	$n = 105$	81.3	55.8	47.4
Stole cash receipts**	62	14.6	40.4	43.6
Stole other assets**	11	4.2	3.8	9.0
Computers used in Scheme:*	47	37.5	19.2	24.4
If yes				
Input data modified*	36	83.3	70.0	73.7
Computer programs modified	3	11.1	0.0	1.3
Output modified after processing	9	38.9	40.0	42.1
Computer system accessed	7	16.7	20.0	10.5
Hid scheme				
No oversight	73	40.9	50.0	40.0
False documentation	50	40.9	26.0	25.3
Poor/false/delayed records	46	18.2	24.0	34.7
Did not respond to survey item	9			
Demographics (mean)				
Dollar losses ($ 000)***	629.9	1,300.8	687.0	259.4
Duration of fraud (mos.)	34.3	41.2	32.5	31.4

*Differences between levels of government are significant at $p = 0.10$.
**Differences between levels of government are significant at $p = 0.05$.
***differences between levels of government are significant at $p = 0.01$.

local = 25.3%) and poor record keeping (federal = 18.2%; state = 24.0%; local = 34.7%) also contributed to the deception.

5.4. Detection of Frauds

The mechanisms through which the reported frauds were detected are summarized in Table 7.

One-fourth of the government frauds reported in this study were ultimately discovered through the audit function (federal = 28.6%; state = 30.2%; local = 25.4%). Outsiders, such as vendors and bank officials, and anonymous hotline tips contributed to initial detection in some cases, particularly at the

Table 7. Characteristics of the Means of Discovery of 178 Cases of Fraud in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Fraud communication channels				
Internal	101	65.9	60.9	62.2
External	55	42.1	45.5	29.7
Mode of discovery				
Internal/external auditors	43	28.6	30.2	25.4
Employee suspicions	77	33.3	48.8	59.2
Outsider suspicions	19	16.7	14.0	8.5
Hotline	14	16.7	7.0	5.6
Suicide/confessed	3	4.8	0.0	1.4
Did not respond to survey item	22			

federal level. Congruent with Albrecht et al. (1995, p. 119) who pointed out "it is co-workers and managers who are in the best position to detect fraud," more fraudulent activity was uncovered through the expression of suspicions by employees (federal = 33.3%; state = 48.8%; local = 59.2%). Given the duration of the frauds, these detection differences underscore the need to pinpoint red flags that will bring such frauds to light in an audit.

5.5. Red Flags Associated with the Misappropriation of Assets

According to respondents, many red flags existed prior to the discovery of the frauds. These red flags are discussed within the categories of risk factors identified in SAS 82. The categories include red flags indicating susceptibility of assets to misappropriation (para. 19a), red flags indicating internal control weaknesses (para. 19b), and personnel conditions conducive to misappropriation (para. 20). Other red flags related to misstatements in financial statements (para. 25), including discrepancies in accounting records and conflicting or missing evidential matter, are also discussed. Problematic or unusual relationships between auditor and client were also mentioned in SAS 82; however, the questionnaire did not address this issue because respondents were unlikely to be aware of such problems.

Responses on the presence of red flags related to the reported government frauds are summarized in Table 8 through 12. Tables 8 and 10–12 list a variety

Table 8. Red Flags Related to Controls That Were Present in 178 Cases of Fraud in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Lack of appropriate management oversight (S)	109	66.7	59.6	59.0
High rate of turnover/management not stable (I)	13	10.4	7.7	5.1
Weak/nonexistent internal controls (S)	96	60.4	50.0	52.6
Exercised unquestioned authority (I)	63	27.1	32.7	42.3
Worked alone/excessive hours	57	29.2	28.8	35.9
Usurped others' job responsibilities (I)	44	20.8	28.8	24.4
Complaints				
Employee/insider	40	20.8	19.2	25.6
Service recipients, vendors, taxpayers	26	12.5	15.4	15.4
Hotline***	16	18.8	3.8	6.4
Internal audit findings ignored	22	16.7	13.5	9.0
Management letter suggestions ignored	12	2.1	5.8	10.3
Other feedback ignored/poorly explained	23	6.7	13.5	10.3
Refused leave (I)	19	8.3	9.6	12.8

Note: (S) Factor specifically identified in SAS 82.
(I) Factor implied in SAS 82.
***Differences between levels of government are significant at $p = 0.01$.

of red flags that were identified by respondents. Some of these red flags represent refinements of more general red flags, such as specific suspicious behavior (e.g. gambling) by the perpetrator or particular discrepancies (e.g. duplicate payments). They are grouped by category of risk factor used in SAS 82. Table 9 summarizes responses regarding key internal control issues related to specific areas where frauds occurred: cash receipts, cash disbursements, payroll (including personnel screening), inventories and fixed assets, and contract administration. The internal control red flags and the failures of specific types of controls are related and are discussed in the following section.

5.5.1. Internal Control Weaknesses

SAS 82 (para. 19b) suggests that specific control weaknesses create a number of red flags. The eight key controls identified are lack of management oversight, lack of job applicant screening procedures, inadequate record keeping, lack of segregation of duties or independent checks, lack of appropriate authorization and approval, poor physical safeguards over assets, lack of timely and appropriate documentation, and lack of mandatory vacations.

Table 9. Internal Control Assessments in 178 Cases of Fraud in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

Internal Control Red Flags Related to Reported Frauds	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Panel A: Adequate controls circumvented				
Cash receipts				
Physical security over receipts	18	4.2	11.5	12.8
Proper documentation	12	4.2	7.7	7.7
Timely creation of records	19	4.2	15.4	1.5
Accurate/complete data entry*	12	0.0	7.7	10.3
Segregation of duties	15	4.2	11.5	9.0
Receipts reconciled/confirmed**	12	0.0	5.8	11.5
Cash disbursements				
Physical security over checks	13	6.3	9.6	6.4
Proper documentation	39	20.8	23.1	21.8
Accurate/complete data entry**	21	20.8	3.8	11.5
Segregation of duties	25	22.9	11.5	10.3
Disbursements reconciled	8	8.5	1.9	3.8
Compare expenditure to appropriation*	6	8.5	1.9	1.3
Monitor encumbrances	4	2.0	1.9	2.0
Payroll/personnel				
Job applicant screening	5	2.1	1.9	3.8
Physical security over checks	4	2.1	0.0	3.8
Proper documentation	6	4.2	1.9	3.8
Accurate/complete data entry	8	6.3	0.0	6.4
Segregation of duties	8	8.3	0.0	5.1
Inventories/fixed assets				
Physical security over assets	4	2.1	1.9	2.6
Proper documentation	4	2.1	1.9	2.6
Accurate/complete data entry	3	4.2	0.0	1.3
Segregation of duties	6	6.3	1.9	2.6
Records reconciled	2	2.1	0.0	1.3
Contract administration				
Bidding process	10	4.2	9.6	3.8
Monitor/evaluate performance*	8	4.2	9.6	1.3
Results against objectives*	8	2.1	9.6	2.6
Panel B: Controls missing				
Cash receipts				
Physical security over receipts***	42	8.3	25.0	32.1
Proper documentation***	12	2.1	15.4	3.8

Table 9. (*Continued*)

Internal Control Red Flags Related to Reported Frauds	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Timely creation of records	28	4.2	11.5	25.6
Accurate/complete data entry***	28	10.4	15.4	19.2
Segregation of duties***	50	8.3	30.8	38.5
Receipts reconciled/confirmed**	44	10.4	32.7	28.2
Cash disbursements				
Physical security over checks**	24	4.2	13.5	19.2
Proper documentation	43	29.2	26.9	19.2
Accurate/complete data entry	23	10.6	15.4	12.8
Segregation of duties	40	20.8	30.8	29.5
Disbursements reconciled	35	18.8	21.2	19.2
Compare expenditure to appropriation	24	12.5	13.5	14.1
Monitor encumbrances	18	14.0	5.6	7.7
Payroll/personnel				
Job applicant screening	7	6.3	5.8	1.3
Physical security over checks	6	4.2	1.9	3.8
Proper documentation	8	4.2	0.0	7.7
Accurate/complete data entry	6	6.3	1.9	2.6
Segregation of duties	8	2.1	3.8	6.4
Inventories/fixed assets				
Physical security over assets	5	4.2	1.9	2.6
Proper documentation	8	8.3	1.9	3.8
Accurate/complete data entry	7	4.2	3.8	3.8
Segregation of duties	6	4.2	1.9	3.8
Records reconciled	8	6.3	3.8	3.8
Contract administration				
Bidding process	9	8.3	7.7	1.3
Monitor/evaluate performance**	16	16.7	9.6	3.8
Results against objectives**	16	16.7	9.6	3.8

*Differences between levels of government are significant at $p = 0.10$.
**Differences between levels of government are significant at $p = 0.05$.
***Differences between levels of government are significant at $p = 0.01$.

This study divides the analysis of key control problems in the particular area of the fraud into two categories: circumvention of controls and lack of controls. With the former, the entity has properly designed controls, but perpetrators are able to circumvent them. Lack of internal control indicates that proper controls either are not in place or are not operating effectively.

As noted in Table 8, respondents reported that internal controls, in general, were weak or non-existent in a majority of cases at each government level (60.4% of the federal frauds, 50.0% of the state frauds, and 52.6% of the local frauds). In their detailed responses, the survey participants actually noted a lack of specific controls 526 times and circumvention of controls was cited 283 times. Detailed indications of control failures are shown in Table 9.

Proper documentation of expenditures (37.1% [39/105] of all cash disbursement frauds) and timely creation of cash receipts records (30.6% [19/62] of all thefts of cash) were the two most frequently circumvented controls. Circumvention of proper controls is particularly troublesome because reliance on controls will be misplaced. Such activities suggest a greater need for targeting red flags outside the internal control system. Two earlier tables may shed further light on this issue.

Lack of management oversight of organizational activities, cited in 109 cases (Table 8), proved to be the most common control deficiency present in the governmental frauds reported herein. This red flag, specified in SAS 82, was observed in 66.7% of the federal frauds, 59.6% of the state frauds, and 59.0% of the local frauds. The respondents also noted a high rate of management turnover or unstable management in 13 cases, which may have contributed to the lack of management oversight. One manifestation of inadequate oversight is management's failure to respond to formal or informal warnings. In a significant number of cases, management disregarded internal audit warnings (federal = 16.7%; state = 13.5%; local = 9.0%), management letter suggestions (federal = 2.1%; state = 5.8%; local = 10.3%), and other feedback that should have alerted the organization that a fraud might exist (federal = 16.7%; state = 13.5%; local = 10.3%). These red flags, none of which are specified in the Standard, take on added significance when one recognizes that fewer than half the organizations even had internal auditors (see Table 4).

As noted in Table 7, over 60% of the victim entities had internal and 30% had external communication channels in place to report fraud. These channels appear to have facilitated discovery of reported frauds. The expression of complaints or concerns by insiders (federal = 20.8%; state = 19.2%; local = 25.6%), by outsiders (federal = 12.5%; state = 15.4%; local = 15.4%), and through hotlines (federal = 18.8%; state = 3.8%; local = 6.4%) may have served as alternatives to formal warning mechanisms. When collusion takes place or controls are circumvented in other ways, fraud communication channels are most important.

Lack of management oversight creates a wide variety of fraud opportunities because such behavior indicates an inability or unwillingness to establish and support internal control systems. Hence, an audit checklist would be more effective if it listed various examples that would call attention to conditions that give rise

to a particular type of fraud. Additional examples of government management oversight failure, shown on Table 11, include excessive budget overruns (19 cases) and excessive budget revisions (14 cases).[4] Failures to monitor contract encumbrances (Tables 9, 8 cases) and compare contract results to objectives (8 cases) are also signs that management is not exercising oversight. Employee, outsider, and hotline complaints regarding improper transactions, which should trigger an alarm that misappropriation may be occurring, may serve as a buffer to lack of response to formal communications, but cannot totally mitigate the problem.

SAS 82 points out that segregation of duties is a fundamental control designed to protect assets from misappropriation. Failure to establish this vital control invites a single employee to use his or her combined responsibilities to derive personal benefits at the expense of the entity because it creates an opportunity to fraudulently obtain assets and hide the fraud. This commonly cited internal control weakness was noted in 104 of the 178 cases (Table 9), and served as a red flag for all types of fraud. Specifically, an inappropriate segregation of functions was present in 72.6% of the cash receipts frauds, 37.1% of the cash disbursements frauds, and 36.4% of the non-cash asset defalcations. This control weakness was observed most often at the local level.

The positive correlation of control weaknesses with the misappropriation of assets illustrates that the perpetrators need power and opportunity to sustain the fraud. As employees gain increased power or access to assets, the organization must devise and maintain additional controls to protect itself from potential fraudulent activities. If an individual seeks out such power or access without adequate organizational reasons, or overrides existing internal controls, concerns should be raised. By intentionally looking for a variety of red flags related to segregation problems, detection is much more likely.

As expected, the study identified additional problem areas in the control system (see Table 9). Many organizations did not provide adequate physical security over cash receipts (62.9% [42/62] of all cash receipts thefts), checks (16.2% [24/105] of the cash disbursements frauds), and non-cash portable assets (36.4% [5/11] of the asset frauds). Untimely or incomplete data entry (40.3%) and inaccurate cash reconciliations (61.3%) proved to be problematic areas associated with cash receipts frauds. Improper documentation contributed to the fraudulent activities noted in the cash disbursements arena (36.2%).

Appropriate contract administration proved to be most difficult to accomplish at the federal level of government. Weakness in monitoring and evaluation of contract performance and failure to compare results against objectives occurred significantly more frequently in frauds at the federal level than at the state or local levels (16.7% for federal vs. 9.6% at the state and 3.8% at the local for both weaknesses). Such weaknesses, coupled with improper documentation,

provide ample opportunities for vendors acting alone or in concert with insiders to implement mutually beneficial fraudulent activities.

SAS 82 (para. 19b) identifies a wide variety of controls designed to prevent workplace situations that may encourage individual fraudulent behavior. The respondents reported several red flags that signaled failures with regard to these controls (Table 8). The perpetrator's exercise of unquestioned authority, suggested as a risk factor in SAS 82, was identified in 63 cases (27.1% of the federal frauds, 32.7% of the state frauds, and 42.3% of the local frauds). Other red flags, not specifically identified in SAS 82 but cited in the survey, included working alone or excessive hours (29.2% of the federal frauds, 28.8% of the state frauds, and 35.9% of the local frauds). Respondents also cited a perpetrator's usurping another employee's job responsibilities (20.8% of the federal frauds, 28.8% of the state frauds, and 24.4% of the local frauds) as a red flag associated with the presence of fraud.[5] While it might be considered an example of lack of segregation of duties, this particular red flag is not otherwise noted in SAS 82. Refusal to take leave is a violation of the need for mandatory vacations noted in the Standard, and it was cited as a red flag in 8.3% of the federal frauds, 9.6% of the state frauds, and 12.8% of the local frauds.[6]

Lack of job applicant screening, a red flag cited in SAS 82, was mentioned in only 7 cases, and it was circumvented in another 5 cases. This and the other results recounted herein suggest that the frequency of red flags will vary in different types of entities and in different areas of the control system. The auditor must be alert to the types of red flags that are relevant and likely to occur in his or her audit area of responsibility. Checklists that target those areas can help assure that key control weaknesses for a particular situation will warn the auditor of potential fraud.

5.5.2. Personnel Conditions Related to Misappropriation of Assets

SAS 82 (para. 20) cites four red flags thought to be useful in identifying situations where employees might be involved in asset misappropriation. These include anticipated future layoffs, employees exhibiting dissatisfaction, observed changes in lifestyle or behavior of employees, and known employee personal financial pressures. Red flags related to personnel conditions cited by respondents are reported in Table 10.

Changing lifestyles (25.0% of the federal frauds, 23.1% of the state frauds, and 34.6% of the local frauds) or experiencing financial difficulties (10.4% of the federal frauds, 21.2% of the state frauds, and 38.5% of the local frauds) both occurred frequently. However, only financial difficulties significantly varied across levels of government – emerging more often at the local level. Other commonly cited personnel problems, not identified in the Standard, included defensive responses or excuses (37 of the cases), preferential treatment of employees or contractors (28 of the cases), and conflicts of interest (25 of the cases). The latter red flag was

Table 10. Red Flags Related to Personnel Conditions That Were Present in 178 Cases of Fraud in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

Red Flags	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Changed lifestyle/increased spending (S)	51	25.0	23.1	34.6
Financial problems (S)***	46	10.4	21.2	38.5
Defensive/excuses	37	16.7	17.3	25.6
Workforce changes				
Decreased (S)	31	28.6	22.4	10.5
Increased**	24	23.8	12.2	10.5
Preferential treatment of employees/contractors	28	20.8	17.3	11.5
Conflict of interest**	25	22.9	5.8	14.1
Tense/irritable***	9	2.1	5.8	9.2
Gambling habit***	6	0.0	0.0	7.7
Alcohol/drugs	5	2.1	1.9	3.8
Lost sleep	2	0.0	1.9	1.3

Note: (S) Factor specifically identified in SAS 82.

** Differences between levels of government are significant at $p = 0.05$.

*** Differences between levels of government are significant at $p = 0.01$.

observed more frequently at the federal and local levels of government. One SAS 82 red flag, general employee discontent, was not noted in the responses.

The organization must be alert to any employee behavior that indicates the possible perpetration of fraud. When an individual spends more money than he or she earns, an occurrence that was cited in 51 cases, it should serve as a warning that fraud may be occurring. Conditions that indicate that the individual is under stress should alert the auditor to potential misappropriation problems. In the frauds reported, financial problems were present in 46 cases, the perpetrator displayed loss of sleep or tension and irritability 11 times, and gambling and alcohol problems were present in 11 cases.

Changes in personnel at the macro level may also serve as a harbinger of fraud. While over half of the defrauded organizations experienced no significant changes in either the number of employees or level of services provided, most changes that did occur were experienced at the federal level. Personnel reductions, cited in SAS 82, occurred in 28.6% of the federal agencies versus 22.4 and 10.5% in state and local governments, respectively. Additional organizational changes, not specified in the Standard, were reported in this study, as well. Personnel increased in 23.8% of the federal agencies versus 12.2 and 10.5% at the state and local levels, respectively.

Increases in services provided were noted in 33.3% of the federal victims, 16.3% of the state entities, and 21.1% of the local governments. Service decreases were reported less frequently (federal = 11.9%; state = 4.1%; local = 1.3%).

These results provide some support for the relationship between organization size changes (personnel or responsibility) and the presence of fraud. The auditor must be sensitive to this possibility. A person whose position is being eliminated might react by perpetrating a fraud as a form of compensation or revenge. Increases in personnel might disrupt the organization's security as it adjusts to new positions and responsibilities. Increases in services provided might create fraud opportunities if the organization is unfamiliar with implications of the service activity.

5.5.3. Discrepancies in Accounting Records

SAS 82 (para. 25) notes that, in addition to the formal list of fraud warnings provided, problems in the accounting records may also provide red flags. These include improper transaction recording, unsupported or unauthorized balances or transactions, and last-minute adjustments. The red flags identified by respondents related to discrepancies in the records are summarized in Table 11.

The most commonly observed red flag in this study related to improper recording of transactions, which tended to be inadequate or incomplete. This red flag is specifically noted in the Standard and affects 33.3% of the federal frauds, 34.6% of the state frauds, and 43.6% of the local frauds. The respondents also noted several red flags not specified, but perhaps implied, in SAS 82 regarding unsupported or unauthorized balances or transactions. These include excessive budget overruns (19 cases), unlocated payees and unusual or unexplained accounts (17 cases each), duplicate payments or unlocated service recipients (15 cases each), and excessive or unjustified cash transactions (13 cases). Unlocated service recipients occurred significantly more frequently at the state level, and excessive or unjustified cash transactions were observed more frequently at the local level. The governmental auditor needs to be aware of the great variety of problems, not separately elaborated, that may flag misappropriations in this area.

The Standard notes last-minute adjustments as a red flag. The respondents identified adjustments that appeared problematic, including excessive budget revisions (14 cases) and unexplained adjustments to receivables or payables (13 cases, which occurred significantly more often at the local level of government). The nature of the adjustment rather than its timing appears to be the strong indication that misappropriation is occurring in government.

An indication not apparent in the Standard is the second-most-common problem with accounting records – excessive cash overruns and shortages. This condition existed in 8.3% of the federal frauds, 3.8% of the state frauds, and 21.8% of the local frauds. The occurrence of cash anomalies was much greater at the local level.

Table 11. Red Flags Related to Discrepancies in the Accounting Records That Were Present in 178 Cases of Fraud in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

Red Flags	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Transaction not recorded properly (S)				
Inadequate/incomplete accounting for transactions (S)	68	33.3	34.6	43.6
Assets sold/appearance of ownership maintained	4	2.1	1.9	2.6
Unsupported or unauthorized balances/transactions (S)				
Excessive budget overruns (I)	19	12.5	9.6	10.3
Unlocated payees (I)	17	10.4	9.6	9.0
Unusual/unexplained accounts (I)	17	8.3	5.8	12.8
Duplicate payments (I)	15	8.3	9.6	7.7
Unlocated service recipients (I)*	15	4.2	15.4	6.4
Excessive/unjustified cash transactions (I)**	13	0.0	7.7	11.5
Payees different names, same address (I)	9	10.4	3.8	2.6
Activation of dormant accounts (I)	6	6.3	1.9	2.6
Excessive voids/credits/returns (I)	5	4.2	0.0	3.8
Excessive second endorsements (I)	4	0.0	3.8	2.6
Excessive change in bad debt write-offs (I)	4	4.2	0.0	2.6
Employees could not be located (I)	4	2.1	3.8	1.3
Assets sold for much less than value (I)	3	4.2	0.0	1.3
Last minute adjustments (S)				
Excessive budget revisions (I)	14	10.4	5.8	7.7
Unexplained adjustments to				
Receivables/payables (I)**	13	6.3	3.8	16.7
Investments/inventory/fixed assets	9	10.4	1.9	3.8
Other accounting record indications				
Excessive cash short/over***	23	5.3	3.8	21.8
Excessive number of bank accounts	8	2.1	1.9	7.7
Increased delinquent receivables	7	4.2	5.8	2.6

Note: (S) Factor specifically identified in SAS 82.
(I) Factor implied in SAS 82.
*Differences between levels of government are significant at $p = 0.10$.
**Differences between levels of government are significant at $p = 0.05$.
***Differences between levels of government are significant at $p = 0.01$.

These findings reinforce the usefulness of several red flags. Failure to properly account for transactions in a given area should raise concern that fraud may exist there. Improper accounting record conditions, accounting treatment that cannot be adequately or reasonably explained, or backed up with proper supporting

documentation (see conflicting or missing evidential matter discussed in the next section), or incorrect or unusual treatments that occur frequently also provide serious warnings of possible fraud.

5.5.4. Conflicting or Missing Evidential Matter

Documentation problems should trigger serious concerns regarding fraud. Documentation in proper form should be demanded for all transactions. Any indication that documents have been destroyed or hidden should result in an investigation, as should evidence of document alteration or reuse. Documentation red flags appeared to be very relevant to the frauds reported in this study. Table 12 offers a summary of these results.

Problems regarding documentation used to sustain the misappropriation were widespread, occurring in 71 cases. The most frequent, which is mentioned in the Standard, is the hiding or destruction of documentary material (22.9% of federal frauds, 32.7% of state frauds, and 47.4% of local frauds). The respondents

Table 12. Red Flags Related to Conflicting or Missing Evidential Matter That Were Present in 178 Cases of Fraud in Federal, State, and Local Governments Reported in a 1998 Survey on Public Sector Fraud.

Red Flags	Total (178)	Level of Government		
		Federal (48) (%)	State (52) (%)	Local (78) (%)
Documentation				
Missing or destroyed (S)***	65	22.9	32.7	47.4
Altered*	65	35.4	25.0	44.9
Used multiple times	26	18.8	13.5	12.8
Unnumbered or non-serially numbered	26	18.8	17.3	10.3
Photocopied (S)	18	14.6	11.5	6.4
Time cards missing or suspicious (S)	13	8.3	7.7	6.4
Reconciliations (S)				
Records to assets delayed/not done	53	27.1	28.8	32.1
Unexplained reconciling items (S)***	29	10.4	7.7	25.6
Subsidiary and control do not reconcile (I)	25	16.7	7.7	16.7
Old outstanding checks cleared	10	4.2	7.7	5.1
General ledger did not balance	6	10.4	5.8	10.3
Missing fixed assets (S)	6	4.2	0.0	5.1
Excessive inventory shortages (S)	4	0.0	0.9	3.8

Note: (S) Factor specifically identified in SAS 82.
(I) Factor implied in SAS 82.
*Differences between levels of government are significant at $p = 0.10$.
***Differences between levels of government are significant at $p = 0.01$.

also noted document alteration, not specified in the Standard (35.4% of federal frauds, 25.0% of state frauds, and 44.9% of local frauds). Both of these warnings were significantly more likely to occur at the local level than at the state or federal level.

Other conditions also existed. Perpetrators used the same documents multiple times to support a transaction (26 cases), unnumbered or non-serially numbered documents (26 cases), and photocopied documents rather than originals (18 cases) to hide frauds. An examination of different areas of the control system indicated that documentation weaknesses arose in 11.3% of the cash receipts frauds, 36.2% of the cash disbursements frauds, and 27.3% of the non-cash asset frauds.

Reconciliation or confirmation problems encompass a final form of conflicting or missing evidential matter cited in SAS 82. Reconciliations and confirmations are performed to provide evidence regarding the completeness and validity of records and transactions. Red flags in this area may warn of misappropriations. Problems with reconciliations indicate that items may not be reconcilable because perpetrators have omitted invalid transactions from the records or have hidden evidence of their fraudulent activities. Failure to perform, or delay of, reconciliations may be a stalling tactic. Items which cannot be adequately explained or that do not make sense should not be accepted by the auditor.

The respondents cited failure to perform reconciliations or confirmations in 88 cases. Three types of red flags related to reconciliations were common. First, reconciliations were often omitted or delayed (27.1% of the federal frauds, 28.8% of the state frauds, and 32.1% of the local frauds). This problem may be implied in the control section of the Standard – either as inadequate record-keeping or lack of independent checks – but is not explicitly identified. In addition, unexplained reconciling items were displayed in 10.4% of the federal frauds, 7.7% of the state frauds, and 25.6% of the local frauds, a significant difference. Finally, subsidiary accounts frequently did not reconcile to control accounts (16.7% of the federal frauds, 7.7% of the state frauds, and 16.7% of the local frauds). Unexplained items are specified in the Standard, and disagreement between control and subsidiary accounts is implied.

Most reconciliation and confirmation weaknesses appeared in the cash receipts cycle (61.3% of the cash receipts frauds) and the inventory cycle (36.4% of the asset frauds). Only 19.0% of the cash disbursement frauds involved failure to reconcile expenditures to records.

6. CONCLUSIONS

This study is an effort to systematically investigate risk factors (red flags) associated with asset misappropriation frauds that occur in government. The results of the

study provide support for applying the risk assessment requirements of SAS 82 in the governmental environment.

We find that theoretical factors cited by auditing literature contribute to identifying fraud in governmental entities. In particular, we identify specific factors cited in SAS 82 associated with government fraud and elaborate on these red flags by providing details of factors that may be related to them. Other red flags that appear more specific to governmental entities have also been identified, thus expanding the total red flags to be considered in the governmental arena.

Finally, the study identifies areas of vulnerability to fraud, both by level of government, and by elements of the control system. Knowledge of such information should enable concerned parties within government to create an organizational atmosphere conducive to: (1) the detection of fraud in a more timely manner; and (2) the creation of an internal control system that would offer greater fraud deterrence.

As with all studies, this research is subject to certain limitations. The survey was limited to gathering information on instances of fraud. Focusing purely on failures, it cannot directly address successful asset misappropriation prevention systems. It, however, may provide useful guidance for improvement. The descriptions of fraud and related controls represent several common types of fraud in government, but they are not an exhaustive set.

This study offers opportunities for extensions. It identifies areas for more detailed examination regarding the continuing problem of fraud and the development of models of fraud prevention and detection in the government sector. It could also facilitate the expansion of models of fraudulent activity that may be unique to governmental entities, such as frauds perpetrated by taxpayers and elected officials.

NOTES

1. SAS 54, which superseded SAS 17, distinguishes illegal acts from the types of fraud that are the focus of this study. Neither is included in the table. The table illustrates the evolution of requirements and additional guidance for audit planning and execution related to fraud that culminated in SAS 82.

2. *U.S. v. Simon*, 425 F.2d 796 (2nd Cir., 1969), *cert. denied*, 397 U.S. 1006 (1970); *U.S. v. Clark*, 360 F. Supp. 936 (S.D.N.Y., 1973); *U.S. v. Natelli*, 527 F.2d 311 (2nd Cir., 1975), *cert. denied*, 425 U.S. 934 (1976).

3. The KPMG *1998 Fraud Survey* reported 437 instances of fraud from a survey population of 5000 organizations, but did not report the response rate. Given that 59% of the respondents cited 5 or more cases of fraud and 41% cited between 1 and 4, the number of respondents could have been no more than 130, or 2.6% of the population. Assuming an average of 2 cases for the 41% and 6 cases for the 59%, still a conservative estimate, the rate of frauds reported would have been 2.0% of the population. Our response rate, to a survey asking much more detailed information about the reported fraud, was 1.4%.

4. The Governmental Accounting Standards Board's Statement 34: *Basic Financial Statements – and Management's Discussion and Analysis – for State and Local Governments* issued in June 1999 will focus more attention on budget overruns and adjustments. GASB 34 requires that state and local governments present the original and the final appropriated budgets for the general fund and major special revenue funds as required supplemental information to the financial statements. It further requires that Management's Discussion and Analysis include analysis of significant variations between original and final budget amounts and between final budget amounts and actual results. Thus, a potential benefit of the new accounting standard is prevention or early detection of fraud.

5. Usurping responsibility often appears in the guise of the "overly helpful" or "extremely conscientious" employee. Such employees may offer to perform duties assigned to other employees, such as processing accounts payable authorizations or reconciling cash collections and deposits, that are intended to serve as controls over the work of the employee who usurps the work, and thereby facilitate asset misappropriation and its concealment.

6. This red flag might receive more internal, as well as external, attention if reporting for compensated absences included analysis of unused vacation days, including the lapse of unused days. Experimentation with service efforts and accomplishments, encouraged by the GASB, could be used to do this.

REFERENCES

Albrecht, W. S., Wernz, G. W., & Williams, T. L. (1995). *Fraud: Bringing light to the dark side of business*. New York, NY: Irwin Professional Publishing.

American Institute of Certified Public Accountants (AICPA) (1939). *Extensions of auditing procedure*. Statement on Auditing Procedure No. 1. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA) (1977). *The independent auditor's responsibility for the detection of errors and irregularities*. Statement on Auditing Standards No. 16. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA) (1988a). *The auditor's responsibility to detect and report errors and irregularities*. Statement on Auditing Standards No. 53. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA) (1988b). *Illegal acts by clients*. Statement on Auditing Standards No. 54. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA) (1997). *Consideration of fraud in a financial statement audit*. Statement on Auditing Standards No. 82. New York, NY: AICPA.

Beasley, M. S., Carcello, J. V., & Hermanson, D. R. (1997). *Fraudulent financial reporting: 1987–1997, An analysis of U.S. public companies*. New York, NY: AICPA.

Committee of Sponsoring Organizations of the Treadway Commission (1992). *Internal control-integrated framework*. New York, NY: AICPA.

Epstein, M. (1993). Accountants and the S&L crisis. *Management Acounting, 76*(February), 24–28.

General Accounting Office (1997). *High-risk series: Medicare*. HR-97–10 (February), GAO.

Green, B. P., & Choi, J. H. (1997). Assessing the risk of management fraud through neural network technology. *Auditing: A Journal of Practice and Theory, 16*(Spring), 1–28.

Loebbecke, J. K., Eining, M. M., & Willingham, J. J. (1989). Auditors' experience with material irregularities: Frequency, nature, and detectability. *Auditing: A Journal of Practice and Theory, 9*(Fall), 1–28.

Mancino, J. (1997). The auditor and fraud. *Journal of Accountancy, 183*(April), 32–36.

Pincus, K. (1989). The efficacy of a red flags questionnaire for assessing the possibility of fraud. *Accounting, Organizations, and Society, 14*(1), 153–163.

Steininger, H. J., Willett, R. B., Donlon, J. J., Stagg, C. A., Goode, L. J., & Carson, J. E. (1997). 1996 CFO survey: Challenges facing federal financial managers in an atmosphere of super-scarcity. *The Government Accountants Journal, 46*(1), 36–39, 43–50.

Thompson, L. H. (1992). *Health insurance: Vulnerable payers lose billions to fraud and abuse*. Report to Chairman, Subcommittee on Human Resources and Intergovernmental Operations, U.S. House of Representatives. Washington, DC: United States General Accounting Office.

National Commission on Fraudulent Financial Reporting (The Treadway Commission) (1987). *Report of the national commission on fraudulent financial reporting*. Washington, DC: U.S. Government Printing Office.

Welch, S. T., Holmes, S. A., & Strawser, J. W. (1997). Fraud in the federal government part II: Characteristics of the schemes – The detection and resolution of the cases. *The Government Accountants Journal, 46*(2), 38–43, 45.

AN EMPIRICAL ANALYSIS OF INTERNAL CONTROL WEAKNESSES UNDER SAS NO. 78: AN EXAMINATION OF STATE AUDIT REPORTS

Edmund J. Boyle, Steven M. Cooper
and Marshall A. Geiger

ABSTRACT

While there has been a considerable amount of research regarding internal control over the past several years, scant empirical research has examined SAS No. 78's integrated five-component depiction of internal control in a government setting. In particular, to our knowledge, no study has assessed the types or frequency of weaknesses under the SAS No. 78 framework using actual internal control system findings. In this study, we examine 32 state department and agency internal control reports to assess how well the theoretical framework captures actual system weaknesses, and to determine the relative distribution of weaknesses across components of the framework.

Our results indicate that the five-component framework was able to effectively classify the 213 reported control weaknesses. Control activities had the highest proportion of identified weaknesses (i.e. around 30%) and monitoring the lowest proportion of weaknesses (i.e. around 10%).

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 183–199

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11009-4

INTRODUCTION

The Committee of Sponsoring Organizations (COSO), an outgrowth of the National Commission on Fraudulent Financial Reporting (Treadway Commission, 1987), has had a considerable impact on how we conceptualize internal control in the U.S. Their report (COSO, 1994) set forth an integrated framework for the evaluation of internal control, which was subsequently adopted by the public accounting profession in SAS No. 78, *Consideration of Internal Control in a Financial Statement Audit* (AICPA, 1995). However, while there is continued interest in internal control evaluations by auditing standard setters (cf., SAS No. 94, *The Effect of Information Technology on the Auditor's Consideration of Internal Control in a Financial Statement Audit*, AICPA, 2001), relatively little research attempts to directly assess the five interrelated internal control components espoused by COSO and codified in SAS No. 78. Our purpose is to examine these theoretical control components using actual internal control system weaknesses. Specifically, we present a descriptive analysis of internal control weaknesses communicated by 32 Rhode Island state agencies categorized according to SAS No. 78. Such an evaluation provides needed feedback on the practical application of internal control currently adopted by the profession, as well as provides information on the frequency of types of internal control weaknesses that are relevant for overall audit planning decisions in government audits.

HISTORY AND CONTENT OF SAS NO. 78

In October of 1986, amid growing concerns about the extent of fraudulent financial reporting, the National Commission on Fraudulent Financial Reporting (i.e. the Treadway Commission) began an extensive study and evaluation of the integrity of our system of financial reporting. In 1987, the Commission issued their final report, which provided numerous recommendations to improve the financial reporting environment and auditing standards, enhance the regulatory and law enforcement environment, and improve education of future participants in the financial reporting process.[1] In response to the Treadway Commission Report, the Committee of Sponsoring Organizations (COSO) developed a comprehensive, integrated model of internal control to offer guidance for creating, adapting and monitoring systems of controls to the risks inherent in financial reporting. This integrated framework was later tailored to practitioners by the Auditing Standards Board (ASB) with the adoption of SAS No. 78 (AICPA, 1995). This new SAS superseded the earlier internal control standard enacted only a few years earlier in SAS No. 55, *Consideration of Internal Control in a Financial Statement Audit* (AICPA, 1988).

The revised framework presents a more sophisticated depiction of internal control as a process designed to provide reasonable assurance regarding the achievement of objectives for reliable financial statements, effective and efficient operations, and compliance with applicable laws and regulations. These broad objectives are directly linked to five interrelated components considered necessary to achieving the internal control objectives. The five components of the framework include the control environment, risk assessment, control activities, information and communication, and monitoring.

As outlined in SAS No. 78, the *control environment* component includes seven related factors: integrity and ethical values, commitment to competence, board of directors/audit committee participation, management philosophy and operating style, organizational structure, assignment of authority and responsibility, and human resource policies and practice. The role and importance of the control environment to the effectiveness of the client's internal control system had been evolving, both stated and implied (see for example, the Cohen Commission (AICPA, 1978a); SAS No. 22 (AICPA, 1978b); and SAS No. 47 (AICPA, 1983)), until it became a formal component of SAS No. 55's three-component internal control structure. SAS No. 78 further refined the meaning of control environment as that component which sets the tone of the organization, by influencing the control consciousness of the entity's employees and establishing the foundation for the remaining components.

In the context of financial reporting, *risk assessment* includes the identification, analysis, and management of risks inherent in the reporting process.[2,3] These risks encompass those events and circumstances which can have adverse effects on the underlying integrity of the management assertions embodied within the financial statements. Risk assessment factors identified by SAS No. 78 focus particularly on changes that can influence the extent of financial statement risks, e.g. changes in the operating environment or corporate structure, or new personnel, information systems, technology, product lines or accounting pronouncements.

The *control activities* component includes policies and procedures designed to ensure that management directives are effectively implemented. Four sub-categories are identified: (1) Performance reviews of actual and budgeted financial information; (2) general and application controls incorporated in both manual and/or automated information processing systems, physical controls over safeguarding of assets, and segregation of duties across critical functions. These broad sub-categories of policies/procedures which can identify breakdowns in control, provide guidance for timely responses, and thus aid in achieving the entity's objectives.

The *information and communication* component has two separate yet integrated factors. The information system incorporates the entity's accounting system,

which consists of the methods and records established to properly account for the organization's transactions, and to maintain accountability for assets and liabilities. Communication, in conjunction with the accounting system, should provide clear guidance of the roles and responsibilities of the organizational personnel.

Monitoring incorporates all management oversight of the organization's systems of internal controls, i.e. management is responsible for establishing adequate networks of control as well as maintaining those systems. Continuous monitoring of internal control performance entails ongoing evaluation of the design of the entity's control model, measuring its effectiveness, and making timely corrections for identified weaknesses.

GOVERNMENT AUDITING STANDARDS

The General Accounting Office (GAO), which serves the Congress of the United States, promulgates and oversees auditing guidelines through the *Government Auditing Standards* (GAS), commonly referred to as the "Yellow Book." GAS articulates generally accepted government audit standards (GAGAS) for both financial audits and performance audits. All government audits, including those which fall under the jurisdiction of the Single Audit Act of 1984, must be performed according to the GAGASs. Essentially, GAGAS standards incorporate the external auditor's generally accepted auditing standards, with additional supplemental general standards for all government audits, and supplemental field work and reporting standards for financial and performance audits.

Congress passed the Single Audit Act of 1984 to promote sound financial management, establish uniform requirements for audits of federal awards, promote efficient and effective use of audit resources, reduce municipal burdens, and ensure that Federal departments/agencies rely upon and use audit work performed under the Act. The Act was later amended in 1996 and revised in 1997 in Circular A-133 *Audits of States, Local Governments, and Non-Profit Organizations* to create more consistency and uniformity among federal agency audits. At the present time, state and local governments are mandated to have a single audit if they receive Federal financial assistance in excess of $300,000, and must report the results of those audits to the appropriate levels of management. The single audit report includes an opinion on the financial statement presentation, a report on relevant internal controls of the financial statements and major programs, and a report on compliance with laws, regulations, and provisions of contracts or grant agreements. Audits which fall under the single audit are more extensive than those performed under GAAS or GAS, in terms of the extent of compliance tests and the detail of the resultant audit report.

PRIOR RESEARCH ON SAS NO. 78 INTERNAL CONTROL COMPONENTS

Despite the considerable impact of the Treadway Commission and COSO on our contemporary conception of internal control, relatively little empirical research has been conducted on their framework for internal control, as currently codified in SAS No. 78. While research on internal control continues to evolve on theory building (e.g. Abdolmohammadi, 1993; Chang et al., 1993; Felix & Niles, 1988; Gadh et al., 1993; Hooks et al., 1994; Houghton, 1991; Kinney et al., 1990; Messier & Austen, 2000; Morton & Felix, 1991; Ponemon, 1994; Smith et al., 2000; Spires, 1991) or guidance for practitioners (e.g. Cashell, 1995; Frazier & Spradling, 1996; Galloway, 1994; Kinney & Felix, 1992; Simmons, 1997; Tanki & Steinberg, 1993), empirical studies are more likely to use internal control as a context to study other research objectives such as framing effects (Emby, 1994; Emby & Finley, 1997), experience (Fredrick, 1991), or information processing (Brown & Solomon, 1990; Chang et al., 1993). Moreover, while some research has emerged regarding the former SAS No. 55 three-component internal control structure, we assess the combined import of the SAS No. 78 components on public, private or governmental systems of control. Nonetheless, some prior studies may be interpreted retrospectively under the new framework in that some prior research has examined some of the SAS No. 78 components to varying degrees.

A significant portion of the extant research performed on specific internal control components concentrates on the control environment. Prior to the enactment of SAS No. 55, research indicates that auditor conception of the client's control environment and its impact on the audit was ambiguous and ill-defined. For example, an early study by Haskins and Henarie (1985) codified a listing of 48 control environment attributes after interviewing experienced auditors, and surveyed 146 auditors from the then Big 8 firms to rank order those attributes. Of the top twenty attributes identified, only nine would be classified as control environment factors under SAS No. 78. Interestingly, follow-up interviews revealed that while the environmental attributes were considered relevant to the quality of the client's internal controls, slightly over fifty percent of the auditors indicated that no alterations would be made to their audit plans as a result of unfavorable control environment conditions. The Haskins and Henarie study highlights the difficulty auditor's face in defining and adapting the audit plan without direct guidance concerning control environment factors. Today, however, the control environment is accepted as a tangible component of a strong system of internal control (Bell & Wright, 1995), which can have a substantive impact on the nature, timing and extent of the audit. Yet, direct research and guidance to link

control environment characteristics to audit procedures is still needed (Pany & Whittington, 2001).

Earlier studies generally document the association between the control environment and the perceived strength/weakness of the entity's control system. In a series of studies, Kreutzfeldt and Wallace (1986, 1990) and Wallace and Kreutzfeldt (1995) used a database of 260 Arthur Andersen audit engagements to examine the relationship between financial statement errors and select internal control structure factors. Additionally, several experiments support the role of the control environment, but also raise some doubts as to the extent of its influence on audit procedures (Dusenbury et al., 1996; Marden et al., 1997; Mayper et al., 1989). More recently Apostolou et al. (2001) found that control environment factors were the most influential on internal and external auditors' assessments of the risk of financial statement fraud.

Asare and Davidson (1995) examined whether financial condition and control procedures influenced auditors' assessments of expected balances of selected accounts. As expected, they found a significant sensitivity to control activities, with smaller predictions of unaudited book value errors as a result of stronger control activities.

Research on the risk assessment component of SAS No. 78 to date has been indirect and has not examined this component as a fundamental and unique element of internal control. Similarly, research on the role of the information and communication component of SAS No. 78 has not been assessed directly in the extant literature. Monitoring has been investigated at length in the internal auditing literature; however, it has received little empirical research attention outside of the internal audit context.

Additionally, while several studies examine systems of internal control in governmental settings in an overall sense, we are not aware of any that examine governmental systems of control under the SAS No. 78 framework. Wallace (1981) applied content analysis to a sample of municipal government reports that included internal control disclosures. She found that inconsistencies in reporting risk assessment, doing cost/benefit analysis, and using diverse reporting formats can influence management's effective use of the report and its sensitivity to internal control weaknesses. Cox and Wichmann (1993) elicited the perceptions of government financial officers on the quality of state and local governmental internal control systems and benefits derived from internal control reports, and Jakubowski (1995) examined the impact of the Single Audit Act of 1984 on the financial management of local governments.

Strand et al. (forthcoming) empirically examined government fraudulent "red flags" using the framework articulated in SAS No. 82, *Consideration of Fraud in a Financial Statement Audit* (AICPA, 1997). Similar to our methodology, they used

an external audit pronouncement to analyze the type and extent of documented frauds in all levels of the government. The authors identified numerous instances of internal control breakdown, which in turn led to fraudulent acts. Moreover, their findings suggest that theoretical audit frameworks offered by the public sector can contribute to our understanding of government auditing issues as well.

Hence, while prior studies have at best analyzed some of the five internal control components individually, to our knowledge no study specifically documents internal control weaknesses according to the entirety of SAS No. 78's integrated component framework, and none use a governmental setting for this analysis.

INTERNAL CONTROL AND THE STATE OF RHODE ISLAND

In 1982, Congress passed the *Federal Managers' Financial Integrity Act*, which mandated that federal agencies establish and evaluate their "internal accounting and administrative controls." Beginning December 31, 1983, all agency directors were expected to report annually on the extent of their agency's compliance with federal guidelines, and also report any material weaknesses identified. Four years later, in response to this federal legislation, the Rhode Island General Assembly formally enacted the Federal Integrity and Accountability Act of 1986. This legislation declared that:

(1) Each public corporation must maintain effective systems of internal accounting and administrative control as an integral part of its management practices.
(2) The systems of internal accounting and administrative control of each state agency shall be evaluated on an ongoing basis and, when detected, weaknesses must be promptly corrected.
(3) All levels of management of the public corporation must be involved in assessing and strengthening the systems of internal accounting and administrative control to minimize fraud, errors, abuse and waste of public and quasi public funds.

In 1987, the Rhode Island State Controller and the Auditor General jointly established a system of reporting and a general framework to guide state agencies in performing evaluations on their systems of internal control. Ultimately, due to budgetary constraints, retirements, and a governmental redirection of organizational goals, the program never achieved the success or improved accountability that had been originally intended by the Federal or State legislation. In 1995, the governor's Director of Administration revived the tabled program and requested the Bureau of Audits (the internal audit agency attached to the Governor's Office)

to collect and review self-assessment internal control reports for fiscal 1996, which were due December 31, 1997. We examine the internal control weaknesses reported for fiscal year ended June 30, 1996 by the Rhode Island state departments and quasi-governmental agencies contained in the state-wide review.

RESEARCH METHOD

Data Collection

All state entities, including both public and quasi-public state agencies, were formally contacted directly after the state's 1996 fiscal year end, and charged with submitting a letter and detailed "self assessment summary" to the governor, due by December 31, 1996. Specifically, the stated purpose for this self-assessment was to "evaluate the strengths, weaknesses, opportunities and threats of goals and objectives." The Bureau of Audits used an open-ended format for reporting agency control system weaknesses to allow agency directors flexibility in communicating their control system observations, and their intended responses to the weaknesses identified in the recently received audit reports.

We requested access to the detailed documentation during the fall of 1998 after the state government work on the reports was essentially completed. By the spring of 1999, we received the necessary approval from the Governors Chief of Staff, the Director of Administration, Auditor General and State Controller. Our study includes all department and agencies that responded to this state-wide assessment.

Data Coding

Each of the authors independently read and classified the control weaknesses identified from the self-assessment reports. Codings were compared and reconciled between two of the authors to generate a preliminary coding set. This set was then compared to the third author's independent coding. Any differences were resolved in arriving at the final coding used for analysis. Table 1 presents examples of some of the system weaknesses categorized under the SAS No. 78 framework.

Individual weaknesses could have been classified into more than one of SAS No. 78's five component categories. As indicated earlier, the authoritative literature states that the internal control components are interrelated. SAS No. 78, for example, encourages auditors to "consider the interrelationships of an entity's control environment, risk assessment, control activities, information and communication and monitoring" (AICPA, 1995, AU319.75) in evaluating the

degree of assurance provided by evidential matter. The Standard also emphasized the importance of gaining a thorough understanding of the substance of the control environment because environment factors "may have a pervasive effect on internal control" (AICPA, 1995, AU319.75), and conversely, "may reduce the effectiveness of other components." (AICPA, 1995, AU319.18). COSO (1994) also addressed the linkages of the five components, and their joint influence on the dynamism of internal control systems. The Committee believed that the control environment served as the foundation for all components. However, they did not limit their discussion of component interrelationships with the control environment. In particular,

> the assessment of risks not only influences the control activities but also may highlight a need to reconsider information and communication needs, or the entity's monitoring activities. Thus, internal control is not a serial process, where one component affects only the next. It is a

Table 1. Examples of Internal Control Weaknesses Categorized into the SAS No. 78 Framework.

Single Component Weaknesses
Control Environment
"Inadequate number of field staff"
"Staff development and training are severely constrained due to budgetary and staff limitations"
Risk Assessment
"Computerization of business office only partially complete"
"Need to evaluate tort claims against the department"
Control Activities
"Lack of segregation of duties in the cash receipt cycle"
"Some divisions reconcile inventories while others do not"
Information and Communication
"The 'Staff Information and Procedures Manual' is obsolete"
"Department needs an automated profiling system to match people with opportunities"
Monitoring
". . . department has not had an independent audit of its central business operations"
"Internal audits are to be conducted"
Multiple Component Weaknesses
Risk Assessment and Information/Communication
"Increases in sales and production of goods and services is over-burdening present resources"
Control Environment and Information/Communication
"A new cost allocation system is needed"
Control Environment and Control Activities
"Limited procedures documentation that provides latitude for non-uniform work product"

> multidirectional interactive process in which almost any component can and will influence another (p. 18).

Accordingly, since the components are not mutually exclusive, a number of weaknesses were classified into more that one component. Table 1 presents some examples of weaknesses categorized into multiple component categories.

The three independent sets of codings were very consistent across authors. Kappa coefficients, representing inter-coder agreement, were 0.89 for the first comparison and 0.91 for the second and final comparison. Both of these coefficients are very strong and indicate a significant ($p < 0.001$) amount of inter-coder agreement.[4] Accordingly, the use of the SAS No. 78 framework allowed the researchers to consistently categorize the internal control weaknesses reported in the governmental audit reports on these agencies. The framework appeared relatively robust with respect to capturing all types of identified control weaknesses into at least one of its component categories.

RESULTS

Table 2 presents the internal control weaknesses identified in the state audit reports. In total, the 32 state audit reports identified 213 internal control weaknesses across the varied departments and agencies audited. Each audit report included mention of at least one internal control weakness in the organization audited. These 213 weaknesses were then classified into the five SAS No. 78 internal control components. Again, since a single weakness could be categorized into more than one component, Table 2 indicates that the 213 weaknesses were categorized 349 times into the five internal control components. The greatest number of components into which any one weakness was classified was four.[5] This, however, occurred only once, along with 13 weaknesses classified into three components. However, the vast majority (over 93%) of the weaknesses were considered to fall into either one or two control components. This array of classifications indicates that most weaknesses were relatively concentrated as to its effect on the internal control systems of the audited organizations. However, there remain a number of instances where the identified weaknesses span a significant portion of the control system as depicted in SAS No. 78.

Table 2 also indicates that the control component containing the most weaknesses was the control activities (CA) component. This one component represented almost 31% (107 out of 349) of all weaknesses identified in the study. The next highest control component was the control environment (CE) component with 23% (81 out of 349) of the identified weaknesses. The component with the

Table 2. Frequency of Internal Control Weaknesses Reported by 32 Rhode Island State Agencies for Fiscal Year 1996 – Categorized by Agency into the Five Internal Control Components of Statement on Auditing Standards No. 78.

Total Number of Weaknesses	Control Environment	Risk Assessment	Control Activities	Information and Communication	Monitoring	Total Number of Control Components[a]
25	9	5	14	10	4	42
25	11	3	15	7	4	40
21	10	6	8	6	3	33
19	7	5	13	4	7	36
13	1	2	8	8	4	23
13	1	2	9	5	1	18
13	4	1	4	7	2	18
9	4	2	4	2	1	13
8	5	3	4	3	0	15
8	1	0	8	1	2	12
8	5	3	1	0	3	12
7	1	4	2	3	0	10
5	4	3	2	1	0	10
4	2	0	1	2	1	6
3	1	0	1	1	2	5
3	1	3	0	0	0	4
3	1	1	0	2	0	4
3	0	0	2	2	1	5
3	1	0	3	1	0	5
3	1	1	2	2	2	8
2	2	0	1	1	0	4
2	2	2	1	0	0	5
2	1	1	1	0	0	3
2	0	2	0	0	0	2
2	1	0	1	2	0	4
1	1	1	0	1	0	3
1	1	1	0	0	0	2
1	1	0	0	0	0	1
1	1	0	0	0	0	1
1	0	1	1	0	0	2
1	1	1	0	0	0	2
1	0	0	1	0	0	1
213	81	53	107	71	37	349

[a] Weaknesses may be categorized into more than one control component.

least number of identified weaknesses was the monitoring (MON) component with roughly 10% (37 out of 349) of the weaknesses falling in this component. A proportions test confirms that the weaknesses were not evenly distributed among the five control components ($\chi^2 = 41.22, p < 0.01$).[6]

In order to test whether categorization of a weakness in any one component was related to simultaneous categorization in any other component, we correlated the five control components across the 213 weaknesses. The only positive correlation is between the control environment and risk assessment components ($r = 0.229$; $p < 0.01$). This correlation is consistent with the interrelated nature of these two components. It also supports the presence of the risk assessment factors embodied in the control environment component of the earlier three-component internal control framework of SAS No. 55. Negative relationships were observed for all the other correlations, indicating that categorization in one component effectively means non-categorization in the other components.

We next assessed whether the types of weaknesses identified by the agency directors were related to the size of the state organization being audited, as measured by total operating budget amount. Correlations between the five control component categories and the total operating budgets indicate a marginally significant ($p = 0.05$) relationship only between the operating budget and the monitoring component. i.e. the larger the auditee, the more frequently a monitoring weakness was observed. All other analyses found no significant relationship between size of the audited state agency/department (based on operating budget) and type of weakness identified.

Finally, we examined the weaknesses identified across the various types of governmental agencies audited. We categorized each audited agency/department into one of the following six categories: General Government (e.g. Department of Administration), Human Services (e.g. Department of Children, Youth and Families), Education (e.g. Department of Elementary and Secondary Education), Public Safety (e.g. Department of Corrections), Natural Resources (e.g. Department of Environment Management), and Quasi-Public Agencies (e.g. Rhode Island Public Transit Authority). We then ran five separate logistic regressions with the individual weakness component the dependent variable (coded 0/1), and the agency type as the independent variable. Since operating budget was found to be somewhat related to type of weakness, we also included total operating budget as an independent variable to control possible effects due to organization size. The results of all five independent analyses indicate that type of governmental agency had no significant ($p < 0.10$) effect on the specific types of internal control weaknesses identified. Thus, the types of internal control weaknesses were consistent across the various types of audited state agencies in this study.

We then assessed whether the six types of audited organizations were related to the raw number of weaknesses identified by the agency directors. In essence, we wanted to test whether different types of organizations exhibited a greater number or fewer internal control weaknesses. Accordingly, we ran an unbalanced ANCOVA using the total number of weaknesses identified as the dependent

variable, the organization type as the grouping variable, and the organization's operating budget as the covariate, in order to control for organization size. The results indicate that the type of governmental agency under audit did not significantly affect ($p < 0.25$) the total number of weaknesses identified by the state agency directors, after controlling for size of the organization being audited.[7]

CONCLUSIONS AND DISCUSSION

We use the SAS No. 78 internal control framework to analyze internal control weaknesses reported in 32 state audit reports. The usefulness of the SAS No. 78 integrated framework is found to be relatively high, in that the researchers were able to independently code the weaknesses into the five-component categories very consistently. While the five-component framework is believed to effectively capture the weaknesses identified, the weaknesses were not evenly distributed among the five components. The control activities component had the highest proportion of identified weaknesses (i.e. around 30%) and the monitoring component was found to contain the lowest proportion of weaknesses (i.e. around 10%). Thus, based on this study, actual control activities performed in the organization remain a very important aspect of the system of internal control, and are most likely to be identified by the auditor as deficient.

Monitoring weaknesses were found to be positively associated with size of the audited organization. The larger the audited agency/department, the higher the likelihood that the audit report would contain a monitoring weakness. Size, however, was not related to any of the other types of weaknesses. Additionally, the type of government agency/department was not related to the number or type of weaknesses identified.

Interpretations of our results may suggest meaningful implications for future research and practice. The predominance of the control activity component highlights the need for government managers to closely review, evaluate and amend their existing network of policies and procedures to insure that they include those designed specifically to prevent or detect control activity weaknesses. The correlation of the monitoring component to size is also significant for practice and research. As government agencies grow in size, the need for monitoring activities increase, as well as the possible need for comprehensive and timely audits of those monitoring mechanisms. In our study, the data suggests that larger government entities should continue to establish monitoring mechanisms as a worthwhile activity of an integrated system of internal control.

Our analysis gives an indication of the robustness of the SAS No. 78 framework for varying types of organizations. Although not originally designed exclusively for

government agencies, based on the results of this study, the framework provided in SAS No. 78 appears useful in evaluating control systems in governmental agencies.

NOTES

1. See Elliott and Jacobson (1987) and McEnroe (1989) for further discussion of the recommendations of the Treadway Commission.
2. The risk assessment component is different from the assessment of audit risk inherent in the audit of financial statements. In an independent audit of financial statements, auditors must initially make assessments of audit risk in order to establish inherent and control risk, and ultimately detection risk.
3. The individual agency directors were not asked to classify their self assessed weaknesses according to a magnitude scale, e.g. reportable conditions vs. material weaknesses. Thus, it is not possible to measure and rate the severity of the weaknesses reported.
4. The percentage of inter-coder agreement across all codings was 96.5 for the first comparison and 97.1 for the second.
5. This weakness related to a computer network installation issue.
6. If we delete the three largest departments from our analyses, all reported results remain substantially unchanged. Accordingly, our results do not appear to be significantly driven by the largest departments.
7. An additional unbalanced ANCOVA using total categories of weaknesses identified produced similar results. No significant relationship was found with type of government agency and total number of control components affected after controlling for size.

ACKNOWLEDGMENTS

We gratefully acknowledge comments from Paul Clikeman, Henry Schwartzbach, Paul Copley, the Editor, and the anonymous reviewers on earlier drafts of this paper. We also thank the Rhode Island state officials that granted us access to their statewide internal control study data.

REFERENCES

Abdolmohammadi, M. J. (1993). Discussion of modeling internal controls and their evaluation. *Auditing: A Journal of Theory and Practice* (Suppl.), 130–134.

American Institute of Certified Public Accountants (1978a). *The commission on auditor's responsibilities: Report, conclusions and recommendations*. New York: American Institute of Certified Public Accountants.

American Institute of Certified Public Accountants (1978b). Planning and supervision. *Statement on Auditing Standards, No. 22*. New York: American Institute of Certified Public Accountants.

American Institute of Certified Public Accountants (1983). Audit risk and materiality in conducting and audit. *Statement on Auditing Standards, No. 47*. New York: American Institute of Certified Public Accountants.

American Institute of Certified Public Accountants (1988). Consideration of internal control in a financial statement audit. *Statement on Auditing Standards, No. 55*. New York: American Institute of Certified Public Accountants.

American Institute of Certified Public Accountants (1995). Consideration of internal control in a financial statement audit. *Statement on Auditing Standards, No. 78*. New York: American Institute of Certified Public Accountants.

American Institute of Certified Public Accountants (1997). Consideration of fraud in a financial statement audit. *Statement on Auditing Standards, No. 82*. New York: American Institute of Certified Public Accountants.

American Institute of Certified Public Accountants (2001). The effect of information technology on an auditor's consideration of internal control in a financial statement audit. *Statement on Auditing Standards, No. 94*. New York: American Institute of Certified Public Accountants.

Apostolou, B. A., Hassell, J. M., Webber, S. A., & Sumners, G. E. (2001). The relative importance of management fraud risk factors. *Behavioral Research in Accounting, 13*, 1–24.

Asare, S. K., & Davidson, R. A. (1995). Expectation of errors in unaudited book values: The effect of control procedures and financial conditions. *Auditing: A Journal of Theory and Practice* (Spring), 1–18.

Bell, T. B., & Wright, A. M. (1995). *Auditing practice, research, and education: A productive collaboration*. New York: American Institute of Certified Public Accountants.

Brown, C. E., & Solomon, I. (1990). Auditor configural information processing in control risk assessment. *Auditing: A Journal of Theory and Practice* (Fall), 17–38.

Cashell, J. D. (1995). The effects of SAS No. 55 on audits of small businesses. *Accounting Horizons* (September), 11–22.

Chang, A., Bailey, A. D., Jr., & Whinston, A. B. (1993). Multi-auditor decision making on internal control system reliability: A default reasoning approach. *Auditing: A Journal of Theory and Practice* (Fall), 1–21.

Committee of Sponsoring Organizations of the Treadway Committee (1994). *Internal control – Integrated framework*. New York: Committee of Sponsoring Organizations of the Treadway Committee.

Cox, C. T., & Wichmann, H., Jr. (1993). The perceived quality of internal control systems and reports for state and local government. *Auditing: A Journal of Theory and Practice* (Fall), 98–107.

Dusenbury, R., Reimers, J. L., & Wheeler, S. (1996). An empirical study of belief-based and probability-based specifications of audit risk. *Auditing: A Journal of Theory and Practice* (Fall), 12–28.

Elliott, R. K., & Jacobson, P. D. (1987). The Treadway Report – Its potential impact. *The CPA Journal* (November), 20–26.

Emby, C. (1994). Framing and presentation mode effects in professional judgment: Auditors internal control judgments and substantive testing decisions. *Auditing: A Journal of Practice and Theory* (Suppl.), 102–115.

Emby, C., & Finley, D. (1997). Debiasing framing effects in auditors' internal control judgments. *Contemporary Accounting Research* (Summer), 55–77.

Felix, W. L., Jr., & Niles, M. S. (1988). Research in internal control evaluation. *Auditing: A Journal of Practice and Theory* (Spring), 43–60.

Frazier, D. R., & Spradling, L. S. (1996). The new SAS no. 78. *The CPA Journal* (May), 40–75.

Fredrick, D. M. (1991). Auditors' representation and retrieval of internal control knowledge. *The Accounting Review* (April), 240–258.

Gadh, V. M., Krishnan, R., & Peters, J. M. (1993). Modeling internal controls and their evaluation. *Auditing: A Journal of Practice and Theory*, 113–129.

Galloway, D. J. (1994). Control models in perspective. *Internal Auditor* (December), 46–52.

Haskins, M. E., & Henarie, R. L. (1985). Attributes and audit impact of client's control environment. *The CPA Journal* (July), 18–26.

Hooks, K. L., Kaplan, S. E., & Schultz, J. J., Jr. (1994). Enhancing communication to assist in fraud prevention and detection. *Auditing: A Journal of Practice and Theory* (Fall), 86–117.

Houghton, C. W. (1991). Inherent risk. *Auditing: A Journal of Theory and Practice* (Spring), 1–21.

Jakubowski, S. (1995). Reporting on the control structures of local government under the single audit act of 1984. *Public Budgeting & Finance* (Summer), 58–71.

Kinney, W. R., Jr., & Felix, W. L., Jr. (1992). Implementing SAS no. 55: An interim report. *Expectation GAP Standards* (May), 86–101.

Kinney, W. R., Maher, M. W., & Wright, D. W. (1990). Assertions-based standards for integrated internal controls. *Accounting Horizons* (December), 1–8.

Kreutzfeldt, R. W., & Wallace, W. A. (1986). Error characteristics in audit populations: Their profile and relationship to environmental factors. *Auditing: A Journal of Theory and Practice* (Fall), 20–43.

Kreutzfeldt, R. W., & Wallace, W. A. (1990). Control risk assessments: Do they relate to errors? *Auditing: A Journal of Theory and Practice* (Suppl.), 1–26.

Marden, R. E., Holstrum, G. L., & Schneider, S. L. (1997). Control environment condition and the interaction between control risk, account type and management's assertions. *Auditing: A Journal of Theory and Practice* (Spring), 51–68.

Mayper, A. G., Doucet, M. S., & Warren, C. S. (1989). Auditors' materiality judgments of internal accounting control weaknesses. *Auditing: A Journal of Practice and Theory* (Fall), 72–86.

McEnroe, J. E. (1989). Treadway commission recommendations: Results at fortune 500 companies. *Internal Auditing* (Fall), 32–35.

Messier, W. F., & Austen, L. A. (2000). Inherent risk and control risk assessments: Evidence on the effect of pervasive and specific risk factors. *Auditing: A Journal of Practice and Theory* (Fall), 119–131.

Morton, J. E., & Felix, W. L., Jr. (1991). A critique of statement on auditing standards no. 55. *Accounting Horizons* (March), 1–10.

National Commission on Fraudulent Financial Reporting (Treadway Commission) (1987). *Report of the national commission on fraudulent financial reporting*. Washington, DC: U.S. Government Printing Office.

Pany, K., & Whittington, R. (2001). Research implications of the current auditing standards board agenda. *Accounting Horizons* (forthcoming).

Ponemon, L. A. (1994). Whistle-blowing as an internal control mechanism: Individual and organizational considerations. *Auditing: A Journal of Practice and Theory* (Fall), 118–130.

Simmons, M. S. (1997). COSO based auditing. *Internal Auditor* (68–73).

Smith, J. R., Tiras, S. L., & Vichitlekarn, S. S. (2000). The interaction between internal control assessment and substantive testing in audits for fraud. *Contemporary Accounting Research* (Fall), 327–356.

Spires, E. E. (1991). Auditors' evaluation of test-of-control strength. *The Accounting Review* (April), 259–276.

Strand, C. A., Welch, S. T., Sharp, F. C., & Holmes, S. A. (forthcoming). An empirical investigation of misappropriation red flags in government: SAS 82 and beyond. *Research in Government and Non Profit Accounting*.

Tanki, F. J., & Steinberg, R. M. (1993). Internal control – integrated framework: A landmark study. *The CPA Journal* (June), 16–20.

Wallace, W. A. (1981). Internal control reporting practices in the Municipal sector. *The Accounting Review* (July), 666–689.

Wallace, W. A., & Kreutzfeldt, R. W. (1995). The relation of inherent and control risks to audit adjustments. *Auditing: A Journal of Accounting, Auditing & Finance* (Summer), 459–481.

A LONGITUDINAL ANALYSIS OF LOCAL GOVERNMENT AUDIT QUALITY

Suzanne H. Lowensohn and Jacqueline L. Reck

ABSTRACT

The quality of governmental audits became an issue when a 1986 General Accounting Office (GAO) study of audits conducted by non-governmental auditors revealed that over one-third were substandard. The results of the GAO study contributed to numerous efforts to improve governmental auditing.

Our study examines desk review statistics for Florida governmental entities for fiscal years 1993 through 2000. Findings indicate that, over time, there has been an overall decline in error rates associated with governmental audits conducted by independent certified public accountants. However, it does appear that a learning curve effect occurs when new requirements are issued. The decline in error rates cannot be attributed to an increase in the proportion of Big 5 firm audits in the Florida governmental audit market during the period studied.

1. INTRODUCTION

Following the highly-publicized financial problems within large city governments in the mid-1970s, widespread attention turned to governmental accountability, resulting in a dramatic increase in the demand for governmental audits. Because

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 201–216

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11010-0

of budgetary constraints, state and federal auditors were unable to perform timely audits of all governmental entities. Thus, many states and federal agencies chose to rely on independent public accountants to meet the increased governmental audit demand (NASACT, 2000; Rubin, 1988).

Given the size of annual governmental expenditures, combined with citizen distrust of government, governmental audit quality was of utmost importance in assuring that public funds were expended for their intended purpose. To determine the extent of substandard audits the General Accounting Office (GAO) conducted a study (U.S. GAO, 1986), which examined a sample of governmental, federal fund audits performed by independent certified public accountants. The study revealed that 34% of the 120 audits examined were not in compliance with auditing standards for the "audit as a whole." These alarming results were believed to stem from non-governmental auditors' lack of understanding concerning the specialized nature of governmental auditing, especially in the areas of compliance with laws and regulations, and appropriate internal control systems. The findings led to increased monitoring of governmental audits, as well as efforts by the GAO and American Institute of Certified Public Accountants (AICPA), among others, to improve the quality of governmental audits performed by independent certified public accountants.

The empirical question of whether governmental audit quality has improved since the 1986 GAO report, given the increased attention by government agencies and the accounting profession, remains unanswered. It appears from the data provided in Brown and Raghunandan (1995) that some improvement in audit quality occurred during the time period 1990–1993; however, many of the efforts made to improve the quality of governmental audits were initiated during and after the time of their study. As a result, Elder (1997) called for further study regarding improvements to governmental audit quality.

In response to Elder's call, we examine the variation in governmental audit error rates, an observable measure of audit quality, in the State of Florida for the eight year time period 1993 through 2000. The results of our study indicate that there was a significant decrease in the overall error rates of governmental audits conducted by independent certified public accountants in Florida during that time.

Our study focuses on the State of Florida since Florida statutes require audits of local governmental entities to be conducted in accordance with government audit standards. Also, the State of Florida has actively implemented AICPA recommendations by evaluating governmental audits, and facilitating education and the exchange of information between local governmental accountants and public accounting firms.

The next section of the paper provides background information on governmental audit quality and governmental audits conducted in the State of Florida. Section 3

describes the data and statistical analysis, while Section 4 discusses the results. The final section addresses implications of the research findings.

2. BACKGROUND

After the GAO issued the report, "CPA Audit Quality: Many Governmental Audits Do Not Comply with Professional Standards" (U.S. GAO, 1986), the accounting profession and governmental agencies promptly responded to the finding of substandard audits (Hepp & Mengel, 1992). In 1987, an AICPA task force published a list of 25 recommendations for the public accounting profession, classified into the following five categories: education, engagement, evaluation, enforcement, and exchange of information (AICPA, 1987). Dubbed the "five E's," the recommendations called for mandatory governmental accounting and auditing training for auditors performing governmental audits, improved audit procurement processes, positive enforcement and peer review activities for audit firms conducting governmental audits, sanctions against substandard audit performance, and better communication between the profession and professional organizations. In the same year, a GAO report (U.S. GAO, 1987) recommended the establishment of an audit procurement system within governmental entities that included four critical attributes (competition, solicitation, technical evaluation, and written agreement) associated with the receipt of a quality audit.

A 1988 revision of *Government Auditing Standards* (U.S. GAO, 1988) addressed the education and evaluation elements of the AICPA recommendations. The revision introduced a requirement that auditors involved in substantial portions of a governmental audit complete 24 of 80 biennial continuing professional education (CPE) hours in subjects directly related to the government environment and governmental auditing. In terms of evaluation, the revision required organizations conducting audits in accordance with government auditing standards to have an appropriate internal quality control system in place. A subsequent revision (U.S. GAO, 1994) strengthened the evaluation element, by requiring an external quality control review at least every three years. The 1994 revision also provided greater guidance regarding internal control issues and testing for compliance, two areas which had proven problematic for independent auditors.

Advances were also made in terms of authoritative guidance related to governmental entities (i.e. exchange of information). In January 1992, Statement of Auditing Standards (SAS) No. 69, *The Meaning of "Fairly Presents" in the Auditor's Report* (AICPA, 1992), designated the Governmental Accounting

Standards Board (GASB) Statements and Interpretations as the most authoritative pronouncements for state and local governmental accounting and financial reporting. Thus, independent auditors who audit governmental entities must express an opinion as to whether the auditee has complied with GASB pronouncements. Entities receiving federal assistance, and their auditors, are subject to *Government Auditing Standards* (U.S. GAO, 1994), which also specify the GASB as a source of generally accepted accounting principles (GAAP).[1] The 1995 issuance of SAS No. 74, *Compliance Auditing Considerations in Audits of Governmental Entities and Recipients of Governmental Financial Assistance* (AICPA, 1995), provided guidance to auditors of governmental entities, especially in terms of compliance with laws and regulations – an area where independent auditors had been particularly deficient.[2]

Although enforcement recommendations have taken the longest to implement, progress has also been made in this area. The AICPA established the governmental technical standards (GTNS) subcommittee of the professional ethics division to investigate allegations of substandard governmental auditing by AICPA members. In severe cases, a trial board can expel a member or take other disciplinary action, including suspension of membership for up to two years, admonishment, imposition of specified CPE requirements, or peer review of the respondent's practice (Hepp & Mengel, 1992). Furthermore, many states have monitoring and enforcement polices (NASACT, 2000); while at the federal level, the inspectors general established debarment procedures for auditors performing substandard audits (Hepp & Mengel, 1992).

In addition to actions taken at the national level emphasizing the "five E's," states work to improve audit quality by implementing the "five E's" through law, regulation and/or guidance. Audit quality at the local level is also affected by the degree to which state law and regulations regarding audit requirements conform to government audit standards (GAS). When state audit requirements are consistent with GAS the complexity of local audits decreases, potentially improving audit quality.

The State of Florida has a high degree of conformance between state requirements and federal audit requirements. Florida adopted *Government Audit Standards* for local governmental audits, and more recently, the state adopted a Single Audit Act (Florida Statutes 215.97), which closely parallels the federal Single Audit Act.

Florida is also active in ensuring quality governmental audits through a variety of other mechanisms. In consultation with the State Board of Accountancy, the Auditor General promulgates rules that specify in detail the scope of local governmental audits and the content of the related audit report. The Auditor General's Office also performs extensive desk reviews of all local government annual audit reports and notifies local governments and the state legislature of

deficiencies. The Auditor General's Office educates local government officials and auditors about significant accounting and auditing issues through seminars, responses to technical inquiries, and participation in externally-sponsored conferences, seminars and local organizational meetings, and continuously provides rule updates and relevant correspondence via its Web-site and mailings. Finally, Florida possesses state-mandated governmental audit procurement practices (Johnson et al., 2003), as suggested by the AICPA's Task Force on the Quality of Audits of Governmental Units (AICPA, 1987).

3. DATA & ANALYSIS

3.1. Data

While empirical evidence associates improved governmental audit quality with educational requirements (Thomas et al., 1998), professional guidance (Deis & Giroux, 1992), active monitoring and enforcement programs (Colbert & O'Keefe, 1995; Wallace & Campbell, 1988), and procurement practices (Copley & Doucet, 1993; O'Keefe & Westort, 1992), no existing study documents improvements in governmental audit quality over time, in light of these factors. Proponents of reform argue that evaluation of governmental audit quality combined with enhanced education requirements, increased exchange of information, engagement guidelines, and GAS reporting requirements improve the quality of governmental audits conducted by public accounting firms.

We use data from the State of Florida, because Florida has adopted each of the "five E's" suggested by the AICPA: (1) Education – Florida's state audit requirements are consistent with *Government Audit Standards*, thus auditors must comply with generally accepted auditing standards and the GAO's educational requirements; (2) Engagement – Florida has audit procurement guidelines within its statutes (Florida Statute 218.391); (3) Evaluation – Florida is one of the few states providing an "active and well funded monitoring program (Wallace & Campbell, 1988)"; (4) Enforcement – results of Auditor General desk reviews are submitted to the Florida State Board of Accountancy for follow-up and potential disciplinary actions; and (5) Exchange information – Florida possesses a strong education and guidance function, evidenced by the aforementioned concerted efforts of the Florida Auditor General.

As with prior studies, we measure audit quality using conformance with professional standards. Specifically, we employ the error rate (nonconformance) of items observed in local governmental audit reports to determine audit quality. We rely on data from 1993 through 2000 to test the hypothesis that

over time there is a significant reduction in the error rates for governmental audits when various recommendations aimed at improving audit quality are adopted.

We also review available information detailing the number of audits conducted by Big 5 audit firms. We do so to ascertain whether any observed changes in audit errors are the result of changes in the proportion of audits conducted by Big 5 firms. This is especially pertinent in Florida, since Hackenbrack, Jensen and Payne (2000) suggest that a statewide bidding restriction (repealed in 1993) allowed large, specialized audit firms to enjoy abnormal profits while providing high quality municipal audits. The 1993 repeal could reduce the presence of Big 5 firms in the Florida governmental audit market.

Data for our study are collected from reports prepared by the Florida Auditor General's Office for fiscal years 1993 through 2000. For each fiscal year, in accordance with Florida Statutes (Section 11.45), the Auditor General conducts a *Review of Local Governmental Entity Audit Reports Prepared by Independent Certified Public Accountants*.

The Auditor General's Office reviews the audits of all counties, municipalities with revenues or expenditures in excess of $100,000, and special districts with revenues or expenditures in excess of $50,000 (State of Florida Office of Auditor General, 2000). The Auditor General's Office conducts two types of desk reviews intended to detect audit errors; a completeness/preliminary review is conducted on all audits submitted, and a comprehensive review is made on a random sample of approximately 60 audits (State of Florida Office of Auditor General, 1994).

A primary purpose of the reviews is to determine if audit reports and financial statements comply with GAS, GAAP and Florida statutes (State of Florida Office of Auditor General, 2001). Each year the Auditor General's Office identifies between 25 and 35 items for review. The items selected vary from year to year, with only a small number of the same items being reviewed each of the eight years included in our study.[3] Additionally, not all items reviewed are relevant to all audits. For example, if one of the items reviewed relates to note disclosures concerning long-term debt and the entity being audited has no long-term debt, the item does not apply to the audit.

In addition to the information on audit report errors, the Auditor General's Office maintains limited data on the characteristics of auditors conducting governmental audits. For the years covered by the study, four years of complete data are provided, which allows for analysis of information concerning the number of audits conducted by Big 5 and non-Big 5 auditors, and the total number of governmental audits conducted by individual audit firms. Information is not available concerning the number of audit errors and the auditor for a given governmental entity.

3.2. Analysis

3.2.1. Measuring the Error Rate

The measure of interest is the error rate for audits conducted by independent auditors in the State of Florida. To help ensure that the number of audit errors is properly weighted, the variable used in the analysis is the percentage of audit errors for each item reviewed by the Auditor General's Office. The audit error percentage is determined by dividing the errors for a given review item by the number of audits to which the item is applicable.

$$\text{Error Rate} = \frac{\text{Errors on Audit Compliance Item Reviewed}}{\text{Number of Audit Reports to Which the Audit Compliance Item Applies}}$$

For example, if the item reviewed is the note disclosures relating to long-term debt, and it is found that 45 audits do not have a note disclosure, and the long-term debt note disclosure is relevant to 750 audits the error rate would be 6%. A percentage is more relevant since using the absolute number of errors would not take into consideration the number of audits to which the item reviewed is applicable. For each of the eight years included in the study, an error rate is calculated for each audit compliance item reviewed by the Auditor General's Office.

3.2.2. Testing Performance

Nonparametric tests are performed to test the hypothesis that audit error rates have declined over time. Nonparametric tests are used since a test of the error rates indicates a non-normal distribution. In addition, the small number of compliance items, on which error rates are based, makes parametric test statistics sensitive to outlier observations. Two tests are used to determine if there is a significant decrease in error rates.

The Cox-Stuart test for trends (Daniel, 1990) is conducted on the overall error rate reported for each year studied. However, the Auditor General's Office did not check for the same audit compliance items each year. To ensure results are not driven by changes in compliance items reviewed, a Page's test for ordered alternatives (Daniel, 1990) is conducted on those compliance items which are held constant over the period studied.[4] As a result, eleven audit compliance items are analyzed for the eight years of the study (see Appendix). Also, because the audits reviewed include municipalities, counties and special districts, analysis is provided separately by entity type to reveal whether entity type is driving the reported results.

The items reviewed by the Auditor General's Office can be divided into three compliance categories – GAS, GAAP (which are reviewed as part of the audit), and Florida statutes (state audit requirements that exceed GAS). We are not interested in the state specific requirements, since the focus of the paper is on recommendations aimed at GAS and GAAP compliance; therefore, Florida compliance items are excluded from our analysis. A separate analysis is conducted on the GAAP items for two reasons: (1) to determine whether GAAP errors are driving the study results; and (2) because of potential differences in the relative importance of the two error categories. Unlike GAS, GAAP is the standard used to measure errors in client accounting systems and is, therefore, directly related to audit quality (Krishnan & Schauer, 2000). Potentially, GAAP-related errors can result in flawed decision-making due to the poor (inaccurate or incomplete) data provided.

4. RESULTS

Table 1, panel A provides the number of audits reviewed each year by the Auditor General's Office. The number of audits reviewed has increased each year, due to increases in the number of cities and special districts subject to audit. Panel B of Table 1 provides descriptive statistics concerning the error rate for all GAS and GAAP items reviewed by the Auditor General's Office. In general, the error rates for independent public accountants in the State of Florida are lower than those reported by Brown and Raghunandan (1995).[5] A review of the data reveals that an anomaly occurs in the 1995–1996 period, when the error rate experiences a fairly large increase before it again declines. A possible explanation for this increase relates to the 1994 GAS revisions and subsequent adoption by Florida. The GAS revisions, applicable for audits of periods ending on or after January 1, 1995, could contribute to errors related specifically to the revisions, or overall errors due to the increased audit complexity. While the adoption of GAS by Florida should decrease audit complexity (reduce errors) in the long-run, confusion regarding implementation and compliance could be expected during the first year of rule changes.

The Cox-Stuart test of trend indicates that for all entities reviewed there is a significant (p-value < 0.05) decline in the GAS and GAAP mean error rates (Table 1, panel B). The decline is attributable to significant declines in the mean error rates for cities (p-value < 0.05) and counties (p-value < 0.01). However, an analysis of the median error rate indicates that only cities experienced a significant (p-value < 0.05) decline in error rates.

As indicated in the previous section, the audit items reviewed are not uniform from year to year. Therefore, using Table 1 results, it is difficult to argue that

Table 1. Volume of City, County and Special District Audits Reviewed by the State of Florida Auditor General's Office Between 1993 and 2000 and Descriptive Statistics for GAS and GAAP Error Rates Detected by These Reviews.

Panel A: Total Number of Audits Reviewed, Displayed by Reporting Entity and Year

Year	Cities	Counties	Special Districts	All Entities
1993	361	66	358	785
1994	366	66	384	816
1995	372	66	393	831
1996	375	66	401	842
1997	376	66	405	847
1998	386	66	414	866
1999	385	66	443	894
2000	382	66	456	904

Panel B: Mean [Median] Error Rates[a] for GAS and GAAP Audit Compliance Items Reviewed, by Reporting Entity and Year

Year	Number of GAS/ GAAP Items	Cities	Counties	Special Districts	All Entities
1993	22	0.142 [0.104]	0.121 [0.078]	0.248 [0.200]	0.176[b] [0.147]
1994	25	0.168 [0.136]	0.129 [0.093]	0.189 [0.143]	0.181 [0.137]
1995	21	0.210 [0.138]	0.141 [0.046]	0.279 [0.222]	0.226 [0.152]
1996	22	0.229 [0.168]	0.175 [0.127]	0.324 [0.321]	0.249 [0.168]
1997	20	0.190 [0.132]	0.119 [0.086]	0.302 [0.261]	0.224 [0.164]
1998	19	0.150 [0.111]	0.104 [0.100]	0.210 [0.171]	0.170 [0.147]
1999	20	0.129 [0.120]	0.094 [0.096]	0.243 [0.200]	0.172 [0.161]
2000	16	0.182 [0.148]	0.107 [0.057]	0.257 [0.203]	0.208 [0.171]
Trend test for significance[b]					
p-Values for means		<0.05	<0.01	n.s.	<0.05
p-values for medians		[< 0.05]	[n.s.]	[n.s.]	[n.s.]

[a] The error rate is calculated as: number of errors detected for an audit review item/number of audits to which the audit review item applied.

[b] *p*-values are provided for the nonparametric Cox-Stuart trend test, and indicate whether there is a significant decrease in error rates over the years 1993–2000. *p*-values test the trend in means [medians]. n.s. indicates not significant.

independent public accountants are increasing compliance with audit requirements since they are evaluated against a different standard each year. As a further test of increasing compliance, we identify eleven audit compliance items that were reviewed in all eight years under analysis, and examine error rates for only these

Table 2. Page's Test of Changes in GAS and GAAP Constant Audit Compliance Error Rates[a] of City, County and Special District Audits (Combined) Reviewed by the State of Florida Auditor General's Office Between 1993 and 2000.

Year	Mean	Median	Minimum	Maximum	Std. Dev.	Sum of Ranks
1993	0.211	0.152	0.054	0.527	0.157	45.50
1994	0.270	0.168	0.100	0.599	0.176	59.00
1995	0.298	0.250	0.097	0.612	0.185	58.50
1996	0.313	0.273	0.080	0.696	0.206	67.50
1997	0.273	0.175	0.071	0.583	0.194	53.50
1998	0.199	0.194	0.058	0.525	0.134	36.00
1999	0.210	0.207	0.052	0.514	0.137	36.50
2000	0.230	0.195	0.050	0.636	0.185	39.50

Page's test for ordered alternatives[b]: Test statistic = 1,900, p-value < 0.05

[a] This analysis includes only those eleven GAS and GAAP compliance items that are reviewed every year of the eight years included in the study.

[b] A one-sided test, indicating whether there is a significant decrease in error rates over time. The test statistic is based on the sum of the ranks.

items. The results are presented in Table 2. Page's test of ordered alternatives indicates that, while the error rate is not monotonically declining, there is an overall significant (p-value < 0.05) decline over the study period. A Cox-Stuart test (not shown) of the trend in error rate for each of the eleven individual items indicates that the decline in error rate is not driven by one or two items. Seven of the items experience significant (p-value < 0.05) declines in error rates. None of the items experience a significant increase in error rates. A listing of the eleven compliance items is provided in the Appendix.

We performed an analysis by reporting entity to determine if the results are being driven by any one type of reporting entity. Table 3 shows that the highest error rates are associated with special districts, while the lowest error rates are attributable to counties. Similar to the overall analysis presented in Table 1, Table 3 indicates that the overall reduction in error rates is primarily driven by cities (p-value < 0.01).[6] Table 3 results include both GAS and GAAP compliance items. Similar results (not shown) are obtained when only GAAP compliance items are considered.

Table 4 provides information concerning the type of audit firm auditing Florida governmental entities. The percentage of governmental audits conducted by Big 5 audit firms has been declining over time,[7] and there has been a fairly steady increase in the percentage of audits conducted by non-Big 5 audit firms that have some experience with governmental entity audits. Therefore, the improving error rates over the eight years studied are not the result of an increasing incidence of Big 5 audits in the Florida governmental audit market.

Table 3. Page's Test of Changes in GAS and GAAP Constant Audit Compliance Error Rates[a] of City (Panel A), County (Panel B) and Special District (Panel C) Audits Reviewed by the State of Florida Auditor General's Office Between 1993 and 2000, Reported by Entity Type.

Year	Mean	Median	Minimum	Maximum	Std. Dev.	Sum of Ranks
Panel A: Cities						
1993	0.178	0.128	0.000	0.439	0.159	44.00
1994	0.230	0.192	0.068	0.547	0.164	62.00
1995	0.284	0.321	0.062	0.643	0.202	67.00
1996	0.293	0.204	0.035	0.776	0.245	60.50
1997	0.240	0.143	0.027	0.639	0.215	47.00
1998	0.185	0.172	0.030	0.485	0.137	42.50
1999	0.152	0.144	0.036	0.271	0.090	36.00
2000	0.203	0.174	0.024	0.679	0.194	37.00
Page's test for ordered alternatives[b]: Test statistic = 1,915, p-value < 0.01						
Panel B: Counties						
1993	0.185	0.136	0.000	0.667	0.194	51.50
1994	0.201	0.093	0.000	0.667	0.240	41.50
1995	0.180	0.055	0.000	1.000	0.309	45.00
1996	0.179	0.120	0.016	0.606	0.190	56.00
1997	0.110	0.062	0.000	0.500	0.152	43.00
1998	0.126	0.104	0.000	0.500	0.148	47.00
1999	0.102	0.101	0.000	0.250	0.086	36.00
2000	0.122	0.047	0.000	0.500	0.161	40.00
Page's test for ordered alternatives[b]: Test statistic = 1,677.5, p-value > 0.05						
Panel C: Special Districts						
1993	0.257	0.209	0.000	0.703	0.200	44.00
1994	0.304	0.215	0.000	0.749	0.246	50.00
1995	0.323	0.246	0.125	0.789	0.208	50.00
1996	0.394	0.392	0.150	0.750	0.204	59.00
1997	0.354	0.318	0.134	0.795	0.199	56.50
1998	0.261	0.262	0.095	0.609	0.169	28.00
1999	0.291	0.241	0.066	0.814	0.229	35.50
2000	0.276	0.215	0.053	0.721	0.234	37.00
Page's test for ordered alternatives[b]: Test statistic = 1,715, p-value < 0.05						

[a] This analysis includes only those GAS and GAAP compliance items that are reviewed every year of the eight years included in the study (eleven items for cities, and ten items for counties and special districts).

[b] A one-sided test, indicating whether there is a significant decrease in error rates over time. The test statistic is based on the sum of the ranks.

Table 4. Descriptive Data on Type of Auditor for Audits Reviewed by the State of Florida Auditor General's Office for City, County and Special District Audits, Select Years.

Annual Percentage of Audits Conducted by Auditor Type			
Year	Big 5 Auditor (%)	Non-Big 5 Auditor[a] ≥ 5 Entity-Type Audits (%)	Non-Big 5 Auditor[b] ≥ 5 Total Audits (%)
1993	21.37	29.91	46.58
1998	13.15	41.79	57.39
1999	11.02	50.06	63.44
2000	10.38	47.66	64.17

[a] Represents non-Big 5 audit firms that conducted five or more entity-type (i.e. city, county, special district) governmental audits during year.
[b] Represents non-Big 5 audit firms that conducted five or more governmental audits during year.

5. DISCUSSION AND CONCLUSION

An earlier study conducted by the General Accounting Office (1986) found an unacceptably high percentage of substandard governmental audits conducted by independent public accountants. Subsequently, governmental auditing standards were modified, and monitoring and education activities have been conducted in an effort to improve the quality of audits performed by independent public accountants. Researchers, such as Deis and Giroux (1992) and Elder (1997), suggest that the efforts made by the various audit standard-setting bodies and oversight organizations should result in improved quality. Our study addresses the question of whether the increased exchange of information, education and evaluation, in conjunction with attention to engagement procurement practices and enforcement, results in improved audit quality. We focus on a single state, Florida, which is active in providing evaluation and enforcement activities and audit procurement guidelines, as well as ongoing education and the exchange of information to independent certified public accountants.

Our results provide evidence that, over time, audit quality improved in the State of Florida. During the time covered by our study (1993–2000) the overall audit error rate significantly (p-value < 0.05) decreased.

An increase in error rates in the years 1995 and 1996 warrants further examination. One possible explanation for increased error rates is the adoption of the revised *Government Auditing Standards* (GAS) (U.S. GAO, 1994) in 1995. These so-called "Yellow-Book" standards exceed generally accepted auditing standards,

issued by the American Institute of Certified Public Accountants, in a number of audit areas; thus, raising the overall complexity of the audit in the short-run. Due to the small sample size of repetitive error items, it is not possible to analyze audit errors year to year to determine if the errors are related to GAS requirements or other factors, such as changes in reporting requirements; however, fluctuations in error rates could be attributable to the introduction of new compliance criteria. For example, Wallace and Campbell (1988) noted problems in Florida audit report compliance when new disclosure requirements were introduced.

Consistent with the suggestions by Hackenbrack, Jensen and Payne (2000), the percentage of Big 5 auditors performing governmental audits in Florida decreased by approximately 50% between 1993 and 2000. It is interesting to note, however, that audit quality did not decline with the decline of Big 5 firms in the Florida governmental audit market.

Our study is limited by the small sample size and limited firm data availability. Additionally, while we hypothesize that the reduction in audit errors is related to evaluation, education and exchange information, a causal relationship has not been established. Measurement error is introduced by the Auditor General's Office in a number of ways, two of which are: (1) a change in the method of applying certain criteria, including judgments of the reviewers regarding materiality; and (2) a change in the method of compiling the list of local governmental entities. Despite limitations, Florida's apparent success in reducing GAAP and GAS-related error rates in governmental audits may have policy implications for other states.

NOTES

1. Although organizations, such as the Municipal and Government Finance Officers Associations (MFOA, GFOA), had previously provided governmental entities and their auditors with guidelines, these organizations had no formal authority to enforce compliance with their suggestions, due in part to state sovereignty, whereby each state determines reporting requirements, if any (Allen & Sanders, 1994). Thus, SAS 69 helped raise the status of GASB issuances.

2. SAS No. 74 supercedes SAS No. 68, "*Compliance Auditing Applicable to Governmental Entities and Other Recipients of Governmental Financial Assistance*," which was issued in 1992.

3. Comparing annual reports over time, it is clear why many items leave the Auditor General's checklist. We observe new review items, with high rates of noncompliance, dwindle to a noncompliance rate of less than 10%, the relative cutoff for removal. Furthermore, the severity of the error items diminished in comparison to prior studies. While the General Accounting Office study (1986) reported omission of financial statements or required reports, many of the items that are defined as errors would not, in and of themselves, deem an audit "substandard."

4. Page's test is similar to Friedman's nonparametric ANOVA; however, Page's test is used when the hypothesis is directional (Daniel, 1990). Additionally, the test is appropriate when samples are related (Daniel, 1990), as occurs in the present study. Page's test uses the sum of ranks obtained from the Friedman's test to construct the test statistic. In the current study, the error rate for each compliance item is ranked by year. The rank assigned to the compliance items is then summed for each year. These sums are used in constructing the test statistic, which allows for determination of a significant order effect. See Daniel (1990) for a more complete description of the Friedman and Page tests.

5. Specifically, in the time period covered by Brown and Raghunandan's study, the lowest error rate was 27%, with the highest rate being 64% (Brown & Raghunandan, 1995, p. 6). Panel B of Table 1 indicates that the highest average error rate in our study is 32.4%, which is related to special districts and occurs in 1996. While differences in the review method used by the Florida Auditor General's Office vs. the method described by Brown and Raghunandan could explain some of the error rate difference, both analyses focus on similar types of audit compliance. The time period covered by Brown and Raghunandan was 1990–1993, while our study covers the time period 1993–2000; thus, there is little overlap of reporting periods.

The later time period covered by our study allows for more complete incorporation of the requirements of the 1994 revision of the *Government Auditing Standards* (U.S. GAO, 1994) and the issuances of SAS No. 69, *The Meaning of 'Fairly Presents' in the Auditor's Report* (AICPA, 1992) and SAS No. 74, *Compliance Auditing Considerations in Audits of Governmental Entities and Recipients of Governmental Financial Assistance* (AICPA, 1995). Elder (1997) indicates that such guidance should be considered in evaluations looking at improvements in government audit quality. Additionally, Brown and Raghunandan (1995) did not analyze, or control for differences such as those related to education, exchange information and evaluation provided by various oversight agencies, such as the state.

6. Notice that the number of compliance items tested for counties and special districts is one less than for cities. This is due to the fact that one of the audit review items does not provide sufficient information to allow for calculation of the individual error rates for the counties and special districts.

7. The reduction in Big 5 audit firms was primarily experienced by cities (from 20.6% of cities audited by Big 5 to only 7.5% of cities audited by Big 5) and special districts (from 20% audited by Big 5 to only 10.1% audited by Big 5). Counties, which experienced lower error rates and no significant decrease in rates, lost only two Big 5 auditors during the same time (from 31.8% of counties audited by Big 5 to only 28.8% of counties audited by Big 5).

ACKNOWLEDGMENTS

We would like to thank Ted Sauerbeck of the Florida Auditor General's Office for the information and data he provided. The comments of the 2002 American Accounting Association's Government and Nonprofit section midyear meeting participants, the editor, and an anonymous reviewer are appreciated.

REFERENCES

Allen, A., & Sanders, G. D. (1994, August). Financial disclosure in U.S. municipalities: Has the governmental accounting standards board made a difference? *Financial Accountability & Management*, *10*(3), 175–193.

American Institute of Certified Public Accountants (AICPA) (1987). *Report of the task force on the quality of audits of governmental units*. New York: AICPA.

American Institute of Certified Public Accountants (AICPA) (1992). *The meaning of 'fairly presents' in the auditor's report*. Statement on Auditing Standards No. 69. New York: AICPA.

American Institute of Certified Public Accountants (AICPA) (1995). *Compliance auditing considerations in audits of governmental entities and recipients of governmental financial assistance*. Statement on Auditing Standards No. 74. New York: AICPA.

Brown, C. D., & Raghunandan, K. (1995, September). Audit quality in audits of federal programs by non-federal auditors. *Accounting Horizons*, *9*(3), 1–10.

Colbert, G. J., & O'Keefe, T. B. (1995, Fall). Compliance with GAAS reporting standards: Evidence from a positive enforcement program. *Auditing: A Journal of Practice and Theory*, *14*(2), 1–16.

Copley, P. A., & Doucet, M. S. (1993, Spring). The impact of competition on the quality of governmental audits. *Auditing: A Journal of Practice and Theory*, *12*(2), 88–98.

Daniel, W. W. (1990). *Applied nonparametric statistics*. Boston: PWS-Kent Publishing Company.

Deis, D. R., & Giroux, G. A. (1992, July). Determinants of audit quality in the public sector. *The Accounting Review*, *67*(3), 462–479.

Elder, R. J. (1997, March). A comment on audit quality in audits of federal programs by non-federal auditors. *Accounting Horizons*, *11*(1), 67–71.

Hackenbrack, K., Jensen, K. L., & Payne, J. L. (2000). The effect of a bidding restriction on the audit services market. *Journal of Accounting Research*, *38*(2), 355–374.

Hepp, G. W., & Mengel, J. F. (1992, June). Improving the quality of government audits. *Journal of Accountancy*, *173*(6), 87–93.

Johnson, L. J., Freeman, R. J., & Davies, S. P. (2003). Local government audit procurement requirements, audit effort, and audit fees. *Research in Accounting Regulation*, *16*, 197–207.

Krishnan, J., & Schauer, P. C. (2000, Fall). The differentiation of quality among auditors: Evidence from the not-for-profit sector. *Auditing: A Journal of Practice and Theory*, *19*(2), 9–25.

National Association of State Auditors, Controllers and Treasurers (NASACT) (2000). *Auditing in the states: A summary*. Lexington, KY: NASACT.

O'Keefe, T. B., & Westort, P. J. (1992). Conformance to GAAS reporting standards in municipal audits and the economics of auditing: The effects of audit firm size, CPA examination performance, and competition. *Research in Accounting Regulation*, *6*, 39–77.

Rubin, M. A. (1988). Contracting for municipal audits. *Association of Government Accountants Journal* (Spring), 25–29.

State of Florida Office of Auditor General (1994). *Review of local government audit reports for the fiscal year ended September 30, 1992*. Report No. 12286. Tallahassee, FL: Office of Auditor General.

State of Florida Office of Auditor General (2000). *Review of local governmental entity audit reports prepared by independent certified public accountants for fiscal year ended September 30, 1998*. Report No. 01–024. Tallahassee, FL: Office of Auditor General.

State of Florida Office of Auditor General (2001). *Review of local governmental entity audit reports prepared by independent certified public accountants for fiscal year ended September 30, 1999*. Report No. 02–017. Tallahassee, FL: Office of Auditor General.

Thomas, C. W., Davis, C. E., & Seaman, S. L. (1998, December). Quality review, continuing professional education, experience and substandard performance: An empirical study. *Accounting Horizons*, *12*(4), 340–362.

U.S. General Accounting Office (U.S. GAO) (1986). *CPA audit quality: Many governmental audits do not comply with professional standards*. AFMD 86-33. Washington, DC: Government Printing Office.

U.S. General Accounting Office (U.S. GAO) (1987). *CPA audit quality: A framework for procuring audit services*. AFMD 87-34. Washington, DC: Government Printing Office.

U.S. General Accounting Office (U.S. GAO) (1988). *Government auditing standards*. Washington, DC: Government Printing Office.

U.S. General Accounting Office (U.S. GAO) (1994). *Government auditing standards*. Washington, DC: Government Printing Office.

Wallace, W. A., & Campbell, R. L. (1988). State boards of accountancy: Quality review and positive enforcement programs. *Research in Accounting Regulation*, *2*, 123–154.

APPENDIX

The Eleven Compliance Items Reviewed Each Year, 1993–2000, by the Florida Auditor General's Office

(1) Disclosures relating to federal awards and reporting requirements.
(2) Disclosures related to the extent of budgetary control.
(3) Disclosures on whether budgetary overexpenditures were contrary to law.
(4) Disclosures on the basis of accounting applied to each budget.
(5) Disclosures on the policy on capitalization of construction interest costs.
(6) Disclosures on fund equity, including plans for liquidating deficit fund balances and the nature and purpose of reserves.
(7) Disclosures concerning the types of investments authorized by legal or contractual provisions.
(8) Disclosures concerning the property tax calendar.
(9) Disclosures on the carrying amounts of investments classified by risk category.
(10) Disclosures and financial statement presentation of pension plans.
(11) Referencing and reporting on internal control and compliance.

AN EMPIRICAL TEST OF PUBLIC CHOICE THEORY: COMPARING UNITED STATES AND UNITED KINGDOM LOCAL GOVERNMENTS☆

Gary Giroux, Andrew J. McLelland and Rowan Jones

ABSTRACT

The Gonzalez and Mehay model is a public choice model successfully used to test the incentives of local government decision makers, including accounting-related control and monitoring. It is a monopoly model that assumes that bureaucrats dominate government decision making, since they have a monopoly position on financial information. A key question is whether this model is generalizable to non-American local governments. This paper compares the use of the Gonzalez and Mehay model for both United States and United Kingdom local governments. British local governments are chosen because information is available and the political structures have interesting similarities and differences to U.S. local governments. Our research seems to be the first to test public choice models: (1) using British governments; and (2) conducting comparative testing across countries. This model works reasonably well for U.K. local districts; however, results differ on some dimensions from U.S. cities. A key point is this model can be compared across countries with

☆An earlier version of this paper was presented at the International Decision Sciences Institute Conference in Athens, July 1999.

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 217–237

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11011-2

reasonable success. Differences suggest that institutional and public policy relationships may differ from country to country, an important consideration for further cross-country analysis.

INTRODUCTION

Public choice models have been used effectively to test the incentive structures of American local governments. The Gonzalez and Mehay (GM, 1985) model was used to evaluate United States (U.S.) local government control systems and information asymmetries (Giroux, 1989; Giroux & Shields, 1993). GM is an analytical monopoly model, which assumes that bureaucrats control government decision making because they dominate accounting information dissemination. More precisely, the bureaucrats know the true costs of goods provided and use this information to maximize their budget (Giroux, 1989). However, Giroux and Shields (1993) demonstrate that certain types of accounting controls, information disclosure, and financial audits limit bureaucratic control for U.S. local governments. This suggests that accounting factors increase government efficiency.

Although the empirical model has been tested only with U.S. local governments, the analytical model developed by GM should be generalizable. The purpose of our study is to test the GM model on comparative samples of United Kingdom (U.K.) and U.S. local governments. Britain is chosen because current economic and demographic data are available and there are interesting similarities and differences between U.S. and British local governments, as well as concomitant accounting and auditing practices.

Public choice theory should be generalizable across countries, but different institutional structures, economic characteristics, and public policy goals exist. Empirical results likely differ across countries. In the mid-1990s both the British and American voters replaced conservative governments with more liberal governmental executives and both experienced economic expansion (approximately three percent growth in gross domestic product annually in both countries). The tax systems are somewhat similar, with heavy reliance on income and payroll taxes, and government spending is based on a combination of national and local decision levels. On the other hand, U.K. is traditionally more liberal, has somewhat higher taxes, has a national health service, and a different relationship between the national and local governments. Different empirical surrogates are likely. Differences in interpreting the GM base model relate to intergovernmental aid (since virtually all U.K. local governmental revenue comes from the national government) and demographic differences (e.g. non-white population is not a major factor in Britain). Because much of the relevant control of local government

rests with the U.K. national government, accounting-related factors are less likely to be significant. For example, auditors are chosen by the Audit Commission, a national agency, not by the local districts. A primary objective of this comparative analysis is to determine to what extent public choice theory is generalizable, when differences such as this exist.

PUBLIC CHOICE OVERVIEW

Public choice models combine economics and political science to evaluate public output decisions based on actor incentives. Key actors include voters, elected officials, bureaucrats, and special interest groups. A major point of public choice is the observation that each individual has unique incentives and strives to maximize his/her utility, even when involving society-wide decisions. Thus, public choice models are similar to agency theory except the focus is on non-market decision making.

In voter behavior politicians are elected to satisfy public welfare based on majority voting rules to aggregate individual preferences (Mueller, 1976). Voters prefer some level of public output produced at minimum cost. Assuming perfect information and high political competition, voters should dominate the process and elected officials should base public spending and tax decisions exclusively on voter preferences (Downs, 1957). Bureaucrats under typical voter behavior models have no choice but to act as agents to maximize public welfare based on voter preferences.

Voters may be relatively uninformed about government decisions, a concept called rational ignorance (Browning & Browning, 1994). That is, the costs of gathering and analyzing information are greater than the perceived benefits. Elected officials have the authority to make all-important financial decisions. Public output is provided by the bureaucracy and voters have limited recourse to bureaucrats.[1] If voters are unhappy with public output levels or taxes, they should vote out the incumbent elected officials. Presumably, elected officials will limit bureaucratic power only if political competition is high; i.e. voters may vote them out of office (Browning & Browning, 1994).

Elected officials have power and benefits associated with public office. It is assumed that the major incentive is the reelection potential (e.g. associated with continued prestige, power, remuneration, perquisites of office). This reelection incentive generally is associated with meeting vocal constituent demands (e.g. special interest groups) rather than direct monitoring of the bureaucracy, partially because of the rational ignorance of the voters. If political competition is low, elected officials have fewer incentives to meet constituent demands (Weingast, 1984).

Bureaucrats are appointed by elected officials and are granted authority by the politicians. Bureaucrats may dominate public output decisions with only moderate checks by voters or elected officials, since they have a monopoly position over financial information and because it is difficult to observe bureaucratic behavior (Bendor et al., 1985). The bargaining power of the bureaucracy is related to their ability to conceal relevant information (Niskanen, 1971). By limiting the financial information available, the bureaucrats can expect to overproduce public goods to enhance their position. This may include shirking, financial slack, and excess consumption (Niskanen, 1971).

This suggests a "bilateral monopoly." Bureaucrats are experts in their areas and have a substantial information advantage. However, elected officials have the authority of government. Elected officials approve budgets and tax levels and control operations through monitoring (including the audit) and penalties (such as firing bureaucrats). Elected officials can use their superior authority to counteract bureaucrats' information monopoly (Bendor et al., 1985; Miller & Moe, 1983). The effect of this interaction of elected officials and bureaucrats is a key concern to public choice modeling.

Control-loss is defined as the cumulative discrepancy between the actions of bureaucrats and desires of voters and elected officials (Tullock, 1965). The two components of control-loss are: (1) the excess supply of public output; and (2) inefficient production (Breton & Wintrobe, 1975). This can be stated in budget terms as the difference in appropriations that the bureaucracy succeeds in obtaining vs. the minimum costs of the preferred budget of the voters (assumed to be determined by the position of the median voter) (Toma & Toma, 1976).

Monitoring and control procedures can be used to limit control-loss and many significant procedures are accounting-related. The budget process is a major control process. Budgets limit spending levels and taxes levied, which require legislative approval. If budgeting is effective, excess spending should be minimized and inefficient production monitored and reduced. Full disclosure using annual financial reports should provide financial information; however, it has not been demonstrated that these disclosures limit the bureaucrats' monopoly position over information.

Bureaucrats may adopt a strategy of subverting both output and shirking by limiting access to information. The use of "budget games" by the bureaucracy when faced with potential budgetary oversight is described in detail in the budgeting literature (e.g. Anthony & Herzlinger, 1975, Chap. 10).

In summary, the key public choice actors have different incentives and it's not clear who dominates the process. In a democracy voter interests should be paramount. However, voters may be apathetic and subject to rational ignorance. Elected officials have the power of office, but may be more concerned with

reelection than administrative control. Bureaucrats are professionals and have monopoly power over financial information, but are subject to legislative approval. It seems likely that relative power will differ for alternative government structures.

THE GONZALEZ AND MEHAY MODEL

The GM model is built on the Niskanen (1971, 1975) bureaucratic framework. It assumes bureaucrats determine public output to maximize their utility, based on their monopoly position over government financial information. Constituents receive a bundle of public output (Q), which is funded by total taxes (T). A balanced budget is assumed, with T equal to total expenditures as a measure of Q. It is assumed that bureaucrats maximize the discretionary budget (somewhat analogous to profit maximization by a monopoly firm), the difference between T and the minimum necessary costs (C[Q]) to produce public output to satisfy voters. The discretionary budget is used by bureaucrats to increase available perquisites and, if possible, increase their incomes.[2]

According to GM, the elasticity of taxation with respect to population depends on the evaluation of the publicness parameter. The model implies that population changes will be matched by proportional spending changes for pure private goods (measured as a population coefficient equal to one), but greater relative spending for quasi- and pure public goods (measured as a population coefficient significantly greater than one).

GM test their theory on California cities using expenditures as a measure of public output and six independent variables: population, population density, mean income, intergovernmental aid, median age, and population change. Positive signs are predicted for population, mean income, intergovernmental aid, and median age. GM predicts that public output rises directly with population if government output behaved as pure private goods. Voters are expected to demand more public output as personal income rose. Because of the "flypaper effect," governmental grants are expected to increase spending rather than replace local revenues. Median age is a measure of dependent population, positively associated with higher spending. If population change has a negative sign, government spending is not keeping up with expanding populations; if a positive sign, spending rises more rapidly than population growth.

Four definitions of public output are used as dependent variables: total spending, police, fire, and parks and recreation. The model proved descriptive using multiple regression, with explanatory power of approximately 80% and the majority of the independent variables significant in the expected direction. Particularly important to their theoretical model, public output is shown to behave as "pure private

goods" with expenditure levels rising proportionately with population (measured with a regression coefficient of population equal to one). Fawson and Giroux (1988) extend GM by testing the model on capital outlays. Their results indicate that capital spending behaves as quasi-public goods (with a population coefficient significantly greater than one).

GM assume no monitoring by voters or elected officials in their analytical modeling; i.e. bureaucratic power depends on their monopoly position over information. The expectation is that public output would be greater than the level desired by the median voter and produced inefficiently, because of bureaucratic shirking and related incentives. Consequently, bureaucratic power is likely overstated in this model. Government regulation mandates annual audits. Budget requirements include control elements and most local governments are subject to balanced budget requirements in some form. The value of monitoring and control is to make camouflaging actual costs of providing government services more difficult; i.e. "as a deterrent against distortion of information" (Breton & Wintrobe, 1975, p. 119). GM do not attempt to measure the moderating effects of monitoring and control or the related interaction of elected officials and bureaucrats.

Accounting controls, financial information dissemination requirements, and financial audits may limit bureaucratic power, resulting in lower public output levels produced with greater efficiency.[3] Giroux (1989) adds a measure of effective auditing to the GM model, which reduces expenditure levels significantly as predicted. Giroux and Shields (1993) add several accounting- and audit-related variables to the GM model and find several variables that reduce spending levels, including unqualified audit opinions and cities awarded a Certificate of Achievement for Excellence in Financial Reporting. However, they also discover the potential for bureaucratic roadblocks to limit the effectiveness of accounting control techniques. These include the use of budgeting surpluses to increase spending levels, higher levels of long-term debt to increase capital outlays, and less than complete financial disclosures.

BRITISH LOCAL GOVERNMENTS

At the start of 1997 there were 472 local governments in the U.K. British local authorities are aggregated for fiscal and control purposes, with revenues and spending regulated by Parliament, although accounting and auditing standards are relatively loose. The U.K. is a unitary state with Parliament the ultimate source of authority. Local governments are subject to the doctrine of *ultra vires*, that limits local governments to do only what the law specifically empowers them to

do. Almost all revenues come from nationally provided sources. Ashworth and Gemmell (1996) suggest that local officials in the U.K. justify overspending by encouraging voters to blame the central government for over-taxation.

Local governments are multi-function organizations that include services such as education, housing, environment, health, personal social services, libraries, parks, museums, buses, roads, and street lighting. They are administered by elected councils and professional staffs. Total spending by local authorities represents about 11% of Gross Domestic Product. Most of the authorities (about 85%) represent from 50,000 to 500,000 people.

Local governments are required to maintain a balanced budget, but deficits can be budgeted as long as they are covered by existing reserves. Limits on long-term borrowing are set by the central government and moneys can be spent only for capital expenditures.

The professional accounting body associated with local governments is the Chartered Institute of Public Finance and Accountancy (CIPFA) and its members are Chartered Public Finance Accountants. All local authorities must be audited annually and auditors are picked by the Audit Commission of Local Authorities in England and Wales, a national government agency. Most auditors are District Auditors that now work for the Audit Commission, although about 30% of the governmental audits are performed by public practice accountants on a contract basis.

The fiscal year of all local governments ends March 31 and financial statements must be published by the following December 31. Fund accounting is used, with each fund having a balance sheet, income and expenditure account, and some have a cash flow statement.

COMPARING U.S. AND U.K. LOCAL GOVERNMENTS

While British government is centralized, there is a constitutional split between the federal government and the 50 states in the U.S. Instead of the fewer than 500 local authorities in Britain, there are over 80,000 American local governments that include cities, counties, school districts, and special districts. U.S. cities are multi-functional and, thus, large cities are in some respects similar to U.K. local authorities. U.K. local authorities are subject to *ultra vires*, but U.S. cities come under the jurisdiction of specific states and are generally fairly autonomous. U.S. cities raise most of their own revenues, with some funding from federal and state sources. Local governments in both countries can borrow long-term from national (and potentially international) capital markets, but these funds are limited to capital expenditures.[4]

Both countries were in the middle of an economic expansion in 1995–1996, the period under study. In the U.K., gross domestic product (GDP) grew at a 2.5% rate, with inflation at 3%. Comparable U.S. data are 3.4% GDP growth and inflation at 3.3%.

Local governments in both countries use fund accounting, formal budgeting procedures, and measure spending using expenditures. Both have professional bodies associated with local government accounting issues: CIPFA in the U.K. and the Government Accounting Standards Board (GASB) in the U.S.

Audits are required annually in both countries. Audit requirements in the U.S. are based on American Institute of Certified Public Accountants (AICPA) standards, General Accounting Office (GAO) regulations and federal laws (beginning with the Single Audit Act of 1984). Most governmental audits are performed by independent CPAs, but state auditors are mandated in some states. Auditors in the U.K. are appointed by the Audit Commission of Local Authorities in England and Wales, mostly district auditors working for the Audit Commission.

Annual reports are issued by local governments in both the U.S. and U.K. In the U.K. the fiscal year ends on March 31 and annual reports issued by the following December 31. Fiscal year-end varies for U.S. governments and annual reports usually issued within six months of the end of the fiscal year.

SAMPLES

The U.K. sample is based on the 64 local districts used by Giroux and Jones (1998). The sample is based on the 1995–1996 fiscal year for English districts and 1994–1995 for Welsh districts. After 1995 Wales moved to unitary districts.[5] Annual reports were requested from the U.K. sample districts and 40 were received. These are used to capture the auditing variables. Since this is substantially less than the full sample, U.K. regression results include both the full and reduced samples.

The U.S. sample is cities over 100,000 in population for fiscal year-ended 1996. We requested annual reports from the 209 U.S. cities over 100,000. We also gathered U.S. Census Bureau data. Complete information is available for 165 cities. We tested for size bias between the responded and non-responding cities and found no significant difference based on a *t*-test. Large cities are used as a reasonable comparison for British districts for two reasons. First, they are about the same population size on average and, second, they are multi-function governments. Also, Giroux and Shields (1993) used large cities (fiscal year 1983) to analyze accounting-related factors related to the GM model.

Table 1. Empirical GM Models: Variables and Sources of Data for U.K. Local Districts and U.S. Cities.

GM Model	U.K. Model	U.S. Model
Dependent variables		
Operating spending	Log of net revenue expenditures (NRE) 1996[a]	Log of current expenditures (TCE) 1996[e]
Capital spending	Log of capital spending 1996 (CS)[c]	Log of capital outlays 1996 (CO)[e]
GM base model		
Population	Log of population (POP) 1996[a]	Log of population (POP) 1996[f]
Population density (population per square mile)	Log of population density (PD) 1996[a]	Log of population density (PD) 1996[f]
Average income	Mean income (INC) 1995[d]	Log of income per capita (IPC) 1989[f]
Intergovernmental aid	Not included	Not included
Population change%	Population change% (PC) 1986–1996[a]	Population change% (PC) 1990–1996[f]
Demographic control	Dependent population% (DP) 1996[b] (under 18 & over 65)	Nonwhite population% (NW) 1996[f]
Accounting related		
Surplus-deficit	Appropriations to (from) reserves (APPROP) 1996[a]	Actual surplus-deficit (ASD) 1996[e] (total government revenues/total operating expenditures)
Long-term debt	Log of total borrowing (DEBT) 1996[c]	Debt per capita (DPC) 1996[e]
Auditor	Public accountant (AUDITOR)[e]	Big six auditor (AUDITOR)[e]
Audit timing	Days to audit report (TIMING)[e]	Days to audit report (TIMING)[e]
	Financial data in £	Financial data in $

[a] Chartered Institute of Public Finance and Accountancy (CIPFA), *Finance and General Statistics*, 1996–1997. Net Revenue Expenditures is a CIPFA definition for operating revenues.
[b] CIPFA, *Personal Social Services Statistics, 1995–1996 Actuals*, 1997.
[c] CIPFA, *Capital Payments, Financing, and Debt Statistics, 1995–1996*, 1997.
[d] Inland Revenue: *Inland Revenue Statistics 1996*.
[e] Annual reports of individual cities, fiscal year-ended 1996 and local districts, fiscal year-ended 1995 or 1996.
[f] U.S. Census Bureau.

EMPIRICAL MODELS

The two empirical models used for the U.K. and U.S. samples are similar to the original GM model with some modifications, which are summarized in Table 1. The

British data are based primarily on the 1995–1996 fiscal year, using information from Chartered Institute of Public Finance and Accountancy (CIPFA) statistics. Two definitions of public output are used: (1) net operating expenditures as a broad measure of current operating spending; and (2) capital spending, both of which are logged because of skewness.

Four variables are almost identical to GM: population, population density (population per square mile), population change (from 1986 to 1996), and average income. Intergovernmental aid was deleted, since in the U.K. virtually all revenue comes from the central government, some of it distributed on a per capita basis. This is quite different from U.S. local governments, where grant requests are generally required for funding, often on a competitive basis. Dependent population is used as a demographic control variable. GM use median age, while Giroux and Shields (1993) use percent of non-white population. Because the demographics in the U.K. differ from U.S. counterparts (e.g. there is a small minority population), percent of population 18 and under and 65 and older is used in the U.K. sample. The prediction is the same as GM: dependent population should increase public output. However, there are additional structural differences. For example, healthcare in the U.K. is provided by the National Health Service. There is an added burden to U.K. public expenditures, but not on local districts. Also, there is a slow trend in the U.K. toward the private sector. For 1995 base case projection of long-term costs for older people is £4.5 billion for local authority expenditures and £4.0 for private expenditures (which is expected to grow more rapidly for the private sector).

Two variables are used to capture financial accounting-related control factors. In the U.K. appropriations to (from) reserves (APPROP) is used to measure surplus-deficit. Local districts are subject to a balanced budget requirement, but deficits can be funded from existing reserves. APPROP measures the difference between total operating expenditures and total revenues. This difference effectively is the net operating surplus or deficit. Under the GM model, bureaucrats seek to maximize the discretionary budget to provide financial slack and other purposes. If the balanced budget works as an effective control, a negative coefficient is expected; i.e. reducing public output. However, if bureaucratic strategy dominates, a positive sign should result. Budget slack provides bureaucrats a method of expanding the discretionary budget through strategic manipulation of budget vs. actual surpluses and deficits (Giroux & Shields, 1993).

Total borrowing is used as a measure of non-revenue spending, which may represent a form of fiscal illusion. Long-term debt allows bureaucrats to increase current spending beyond immediate revenues available. The rationale is that voters may be unaware of spending from non-tax sources or that the debt must be repaid with interest (Wagner, 1976). A negative sign is associated with effective control (that is, the disclosure of total debt should moderate further capital outlays

if constituents understand the relationship of new debt to increased future taxes (Giroux & Shields, 1993)), a positive sign with bureaucratic strategy.

Two variables test the impact of the financial audit on government spending. The value of the audit is "as a deterrent against distortion of information" (Breton & Wintrobe, 1975, p. 199). Thus, the audit should have a moderating effect on bureaucratic behavior. In the U.S. high quality audits are associated with using Big Six accounting firms (De Angelo, 1981). A Big Six dummy variable is used for testing. Following Giroux (1989) cities using high quality auditors should have lower spending levels. Most British government audits are conducted by District Auditors. However, about 30% are private accountants selected by the Audit Commission. It is not known if they provide higher quality audits and no sign is predicted.[6] The second variable is audit timing, the number of days from the end of the fiscal year to the audit report date. Early report dates are considered "good news," a combination of an efficient audit and few accounting difficulties to be resolved before the audit report is submitted (Dwyer & Wilson, 1989). The relation to spending levels is not clear. A positive sign indicates that longer audit timing results in greater spending, probably associated with increased financial complexity. A negative sign would indicate higher efficiency associated with reduced spending.

Accounting variables are significant in earlier tests by Giroux (1989) and Giroux and Shields (1993), based on U.S. cities. However, it is less likely that they will be effective for U.K. districts, since most of the financial process is controlled by the central government. Bureaucrats are more likely to dominate the budget process when control is split between central and local elected officials and local officials have limited ability or incentives to reduce tax levels. Studies by Jones and Pendlebury (1991) and Pendlebury and Jones (1985) indicate marked non-compliance of U.K. local districts with accounting requirements accompanied by clean audit opinions and little if any public comment. It is not clear if standard monitoring and control techniques are useful for promoting more efficient government spending under these circumstances.

DESCRIPTIVE ANALYSIS

Panel A of Table 2 presents descriptive statistics for the U.K. sample. Capital spending, on average, is over 10% of net revenue expenditures, and both spending measures have a large standard deviation. Population (POP) ranges from 27,000 to over one million. Large urban areas such as London are divided into separate boroughs. Although the U.K. has the highest population density in Europe, some Welsh counties have the lowest; consequently, PD has a large range. Population change averages a small 6,300 (2.6%) increase, but the range is from −67,000

Table 2. Descriptive Statistics for Samples of 64 U.K. Local Districts for Fiscal Year-Ended 1995 or 1996 and 165 U.S. Cities With a Population in Excess of 100,000 for Fiscal Year-Ended 1996.

	Mean/ Frequency	Standard Deviation/ Percentage	Minimum	Maximum
Panel A: U.K. data				
Dependent variables				
Net revenue expenditures (NRE) (£000)	149,024	166,374	4,039	579,036
Capital spending (CS) (£000)	16,369	19,587	0	104,624
GM base				
Population (POP)	243,675	218,059	27,100	1,035,500
Population density (PD)	12.6	17.8	0.2	90.2
Mean income (INC) (£)	14,625	3,328.9	10,300	27,900
Population change (PC)	6,336	15,906	−66,600	54,300
Dependent population (DP) (%)	60.6	0.21	57.6	70.4
Accounting				
Appropriations (APPROP) (£000)	−2,789	2,784	−11,818	1,920
Total borrowing (DEBT) (£000)	106,294	159,739	0	801,893
Audit timing (TIMING) (Days)	247.7	39.4	157.0	300.0
Public accountant (AUDITOR)	13	33%		
Panel B: U.S. data				
Dependent variables				
Current expenditures (TCE) ($000)	345,205	561,475	44,788	4,026,960
Capital outlays (CO) ($000)	41,104	56,549	166	334,829
GM base				
Population (POP)	312,661	410,778	100,000	3,554,000
Population density (PD)	3.9	2.6	0.15	15.7
Income per capita (IPC) ($)	14,246	3,329	6,284	27,092
Population change (PC)	12,582	30,095	−108,000	175,000
Nonwhite population (NW) (%)	28.6	16.1	2.0	82.7
Accounting				
Actual surplus-deficit (ASD)	1.03	0.086	0.69	1.38
Debt per capita (DPC) ($)	903.7	626.2	80.3	4,357.6
Audit timing (TIMING) (Days)	125.1	36.1	57.0	229.0
Big six (AUDITOR)	88	53%		

to +54,000. The dependent population averages over 60%, with relatively little variation. The extent to which current spending exceeds revenues (APPROP) indicates substantial deficit spending (averaging 2% of revenue expenditures). Given the balanced budget requirement, this is somewhat surprising. Thirteen (33%) of

the audits are conducted by private accounting firms, rather than district auditors. All are Big Six firms. On average, the time to issue an audit report is 248 days (over eight months).

Panel B of Table 2 presents descriptive data for U.S. cities. The spending relationships are similar to U.K. districts. Capital spending is over 10% of operating spending and standard deviations are large. Population is slightly larger than U.K. cities (because of sample selection) and population density considerably lower. Population change is somewhat higher and with a larger range. Average income (stated in dollars) is considerably lower than in Britain. Non-white population is under 30%. On average, U.S. cities have a 3% surplus, rather than the substantial deficits of U.K. districts. Debt per capita is less than $1,000 (the comparable figure for U.K. districts is £431, about $700). Just over half the cities (53%) are audited by Big Six firms. It takes 125 days (over four months) to issue an audit report, about half the time for British governments.

Correlation matrixes are presented in Table 3. Correlations of dependent variables to independent variables suggest expected relationships. Thus, population is highly correlated to all definitions of spending. A severe multicollinearity problem is present between POP and governmental grants with the U.K., with a correlation above 0.9. Because of this problem (which is corroborated by other tests), grants were eliminated from the model. Several other independent variables are significantly related to spending definitions. No other multicollinearity problems across independent variables are detected.

REGRESSION RESULTS

Regression results are summarized in Tables 4 and 5. Table 4 presents the U.K. models and Table 5 the U.S. results. The GM model is a log model. However, several variables have zero or negative values in the current samples. Consequently, a semi-log model is used for this analysis. Logs are used for the dependent, money-denominated, and skewed variables. Both full and reduced U.K. models are run, where the reduced models include the audit variables.

Regression diagnostics are performed for all regression models. These include an analysis of correlation coefficients and variance inflation factors (multicollinearity); Glejser Test and residual plots (heteroscedasticity); stem and leaf and box plots (normality of residuals); studentized residuals and Cook's D (extreme values). As previously stated, multicollinearity is detected between POP and government grants and grants is eliminated from the model. Heteroscedasticity is detected for the full U.K. NRE model. White's correction is run. The correction does not significantly change the *t*-values for the independent variables

Table 3. Descriptive Statistics (Pearson's Correlations) for Samples of 64 U.K. Local Districts for Fiscal Years-Ended 1995 or 1996 and 165 U.S. Cities With Populations in Excess of 100,000 for Fiscal Year-ended 1996.

Panel A: U.K. Model ($n = 64$)[a]

	CS	POP	PD	INC	PC	DP	APPROP	DEBT	TIMING	AUDITOR
NRE	0.720*	0.942*	0.182	−0.018	0.163	0.114	−0.586*	0.524*	−0.274***	0.148
CS		0.995*	0.311***	0.016	0.163	0.223***	−0.411**	0.798*	−0.234	0.252
POP			0.022	0.009	0.347**	0.035	−0.569*	0.422**	−0.272***	0.110
PD				0.310***	−0.085	0.596*	−0.092	0.386**	−0.182	0.279
INC					−0.129	0.692*	−0.051	−0.089	−0.321***	0.325***
PC						−0.140	−0.103	0.019	0.089	−0.235
DP							−0.062	0.220***	−0.329***	−0.385***
APPROP								−0.104	0.395***	−0.193
DEBT									−0.026	0.156
TIMING										0.485**

Panel B: U.S. model ($n = 165$)[b]

	CO	POP	PD	IPC	PC	NW	ASD	DPC	TIMING	AUDITOR
TCE	0.827*	0.902*	0.472*	−0.005	−0.472*	0.394*	−0.088	0.302*	0.274**	0.168***
CO		0.774*	0.336*	0.093	0.083	0.280**	−0.107	0.333*	0.189***	0.129
POP			0.287**	−0.026	0.188***	0.297*	0.002	0.168***	0.203**	0.149***
PD				−0.029	−0.373*	0.388*	−0.174***	0.168***	0.223***	0.081
IPC					0.107	−0.244**	0.095	0.021	−0.307*	0.101
PC						−0.235**	0.210**	−0.057	−0.133***	0.100
NW							−0.267**	0.320*	0.233**	0.134
ASD								−0.129	0.073	0.004
DPC									0.152***	0.120
TIMING										0.062

Note: TIMING and AUDITOR for U.K. districts are based on $n = 40$.

[a]Where: NRE = Net Revenue Expenditures (£000), CS = Capital Spending (£000), POP = Population, PD = Population Density (population per square mile), INC = Mean Income (£), PC = Population Change%, 1986–1996, DP = Dependent Population (% of population under 18 and over 65), APPROP = Appropriations to (from) Reserves 1996, DEBT = Total Borrowing for 1996, TIMING = Days to Audit Report, and Auditor = 1 if Public Accountant Used.

[b]Where: TCE = Current Expenditures for 1996, CO = Capital Outlays for 1996, POP = Populationf for 1996, PD = Population Density (population per square mile) for 1996, IPC = Income Per Capita for 1989, PC = Population Change%, 1990–1996, NW = Non-white Population% for 1996, ASD = Actual Surplus/Deficit for 1996, DPC = Debt Per Capita for 1996, TIMING = Days to Audit Report, AUDITOR = 1 if Big Six Auditor

*Significant at 0.0001.

**Significant at 0.01.

***Significant at 0.1.

and, therefore, is not presented. Extreme values are found in both U.S. models. Table 5 presents the results after removing these observations. Heteroscedasticity also is found. White's correction does not significantly chance the *t*-values and therefore not presented. Nine U.K. districts which have no capital spending for the year (four in the reduced model) are deleted from the capital spending model.

Table 4. Empirical Results Using OLS Regression – U.K. Local District Data for Fiscal Year-Ended 1995 or 1996 Coefficients (*t*-Values).

GM Model	Predicted Sign	Dependent Variables			
		Log of Net Revenue Expenditures (NRE)		Log of Capital Spending (CS)	
		Full	Reducd	Full	Reduced
Intercept		22.4	22.2	18.5	13.5
GM base					
Population (POP)	+	1.61 (14.24)*	1.62 (2.91)*	0.92 (6.43)*	1.07 (5.52)*
Population density (PD)	?	0.05 (0.84)	0.16 (2.09)***	0.11 (1.60)	0.20 (2.39)***
Mean income (INC)	+	−1.66 (−3.41)**	−1.55 (−2.63)***	−1.30 (−2.11)***	−0.84 (−1.18)
Population change (PC)	?	−0.01 (−1.69)***	−0.01 (−1.45)	−0.00 (−0.22)	0.00 (0.04)
Dependent population (DP)	+	−0.10 (−2.06)***	−0.013 (−2.18)***	−0.06 (−1.04)	−0.06 (−0.87)
Accounting related					
Appropriations (APPROP)	+/−	−0.00 (−2.29)***	−0.00 (−1.02)	−0.00 (−1.49)	−0.00 (−0.28)
Total borrowing (DEBT)	+/−			0.05 (1.59)	0.00 (1.29)
Audit timing (TIMING)	+/−		0.00 (−0.48)		0.00 (0.07)
Public accountant (AUDITOR)	+/−		−0.11 (−0.48)		0.14 (0.60)
F-value		76.80*	41.53*	23.48*	21.18*
Adjusted R^2		0.878	0.893	0.745	0.838
Sample size		64	40	55	36

*Significant at 0.0001.
**Significant at 0.01.
***Significant at 0.1.

Table 5. Empirical Results Using OLS Regression – U.S. Cities With Populations in Excess of 100,000 for Fiscal Year-Ended 1996 Coefficients (*t*-Values).

GM Model	Predicted Sign	Dependent Variables	
		Log of Total Current Expenditures (TCE)	Log of Capital Outlays (CO)
Intercept		5.9	3.8
GM base			
Population (POP)	+	1.21 (17.39)*	1.28 (9.40)*
Population density (PD)	?	0.02 (0.39)	−0.10 (−1.11)
Income per capita (IPC)	+	0.00 (2.98)**	0.69 (2.77)**
Population change (PC)	?	−0.01 (−4.63)*	−0.00 (−1.42)
Non-white population (NW)	+	−0.00 (−0.36)	−0.00 (−1.69)***
Accounting related			
Actual surplus-deficit (ASD)	+/−	−0.86 (−2.57)***	−1.19 (−1.83)***
Debt per capita (DPC)	+/−		0.00 (4.16)*
Audit timing (TIMING)	+/−	0.00 (1.92)***	−0.00 (−0.95)
Big six (AUDITOR)	+/−	0.17 (2.80)**	0.10 (0.86)
F-value		132.64*	31.81*
Adjusted R^2		0.867	0.640
Sample size		163	157

*Significant at 0.0001.
**Significant at 0.01.
***Significant at 0.1.

The first model in Table 4 tests the net revenue expenditures full model using the U.K. sample and has an adjusted R^2 of 88%, significant at 0.0001. Five of six GM base model independent variables are significant. The coefficient for POP at 1.6 indicates that public spending behaves as a quasi-public good. This is much higher than the population coefficient for U.S. governments. The negative coefficients for mean income is surprising, indicating that relative spending declines as incomes increase. Public choice theory predicts that public spending rises with income levels (which is the finding with the U.S. sample). However, much of the revenue from the central government is distributed on the basis of a regression model driven by quantified "need." Therefore, the more demonstrated public need (usually associated with lower incomes), the larger the relative central government grant. Population change has a negative sign, suggesting that spending doesn't keep up with changing population (in both directions), consistent with U.S. results. Dependent population (under 18 and over 65) has an unexpected negative sign and is significant. This suggests that spending declines as the percent of dependent population rises.[7]

The negative sign for APPROP indicates that spending is higher as the deficit increases, which is expected if control rather than bureaucrats dominates the process. The U.S. model has the same result. The implication is that this is an effective control mechanism. Giroux and Shields (1993) have negative signs on similar variables, associated with strategic behavior of bureaucrats (e.g. building financial slack in the original budget).

The reduced NRE model has similar results. Adjusted R^2 is 89% and four of five base model variables are significant. However, population density is significant and population change is not. None of the accounting-related variables are significant. This is disappointing, since no evidence is provided for the effectiveness for government spending efficiency. On the other hand, it is not surprising since Parliament maintains financial controls of the local districts, which has fewer incentives than local elected officials to contain the local bureaucracy.[8]

The capital spending full model has an adjusted R^2 of 75%, also significant at 0.0001. However, only two independent variables are significant, population and mean income. Mean income has an unexpected negative sign (as with the revenues expenditures model). The accounting variables are not significant. The reduced model results are similar. Adjusted R^2 is 84% and only two base model variables (population and population density) are significant. Consequently, the GM model is not descriptive of capital outlays in the U.K. (the U.S. model in contrast is more descriptive).

The first model in Table 5 tests total current expenditures for U.S. cities. The model has an adjusted R^2 of 87%, significant at 0.0001. Three of five independent base model variables are significant, all with expected signs. The negative sign for population change is similar to U.K. model results, indicating operating spending lags changes in population. The three accounting-related variables are significant. Actual surplus-deficit is negative and significant, consistent with the similar variable for U.K. districts, but opposite Giroux and Shields (1993) for 1983 data. The positive sign for audit timing suggests greater financial complexity related to higher spending levels. A Big Six audit is associated with higher current spending, the same result as Giroux and Shields (1993).

The capital outlays model has an adjusted R^2 of 64%, significant at 0.0001, and five of nine independent variables are significant. Non-white population has an unexpected negative sign, similar to the dependent population variable in the U.K. sample. This suggests a consistent conflict with public choice theory. A possible explanation is the shift since 1980 to conservative governments in both countries. Welfare spending is reduced for governments investigated.[9] Actual surplus-deficit is negative and significant, similar to the U.K. net revenue expenditures results, suggesting that control dominates the spending process and opposite Giroux and Shields (1993). Debt per capita is positive and significant, suggesting that

bureaucrats dominate the process; i.e. increasing public output by borrowing long-term. This result is the same as Giroux and Shields (1993). Neither audit variable is significant.

The U.S. results are similar to previous studies, such as Giroux and Shields (1993). Here the model is descriptive and provides some support for the importance of accounting practices in limiting government spending levels based on control and monitoring functions. The results for U.K. local districts are mixed. The base model is reasonably descriptive; however, mean income and dependent population have unexpected negative signs. These can be explained, in part, by unique circumstances in Britain. This suggests that institutions and public choice relationships differ across the two countries. The only significant U.K. accounting-related variable is appropriations in the full NRE model, suggesting that accounting (at least as tested here) may not provide effective controls. This is not unexpected, since the central government provides resources and directions for spending funds.[10]

CONCLUSIONS

The Gonzalez and Mehay (1985) model of public choice works modestly well in describing the level of both British and American public expenditures. Results are somewhat similar across the two countries. A key point is that this model can be compared across countries with reasonable success. However, there are differences across the models and also when compared to the earlier study of Giroux and Shields (1993). Differences suggest that public policy relationships may differ across countries, an important consideration for further cross-country analyses. These differences are not well understood. There are structural differences between the U.S. and U.K. and incentives may differ between elected officials and bureaucrats across the countries. The role of Parliament for British local districts is much different (e.g. they exert more direct control) than the role of the federal and state governments over U.S. cities. The development of public choice empirical models and the effectiveness of accounting and auditing controls may differ somewhat across countries.

There are substantial limitations to this study. First is the model comparison across countries. The U.K. model differs from both the original GM model and the current U.S. model. Although the models are comparative, there are no direct statistical tests of differences. The local governments are not identical and the model development differs, suggesting possible difficulties with statistical comparisons. The tests represent limited samples and only a single point in time. There may be structural changes over time that have policy implications. For example, % of non-white population is positive and significant in Giroux and Shields (1993)

based on 1983 data, but negative in the current analysis. Surplus/deficit variables are consistently negative for both U.K. and U.S. samples, but positive in Giroux and Shields (1993). Finally, only limited testing of control and monitoring factors has been conducted.

NOTES

1. Public output is defined as the goods and service provided by a government unit, usually measured by total expenditures. This is sometimes called public goods, but in the GM framework "publicness" is measured from "pure public goods" to "pure private goods," associated with the coefficient of population rather than a measure of expenditure levels.
2. Niskanen (1971) predicts that the utility of the bureaucrat is based on maximizing the budget. Migue and Belanger (1974) assume that the bureaucrat is concerned with managerial discretion and it's the discretionary budget being maximized. The GM model is based on the Migue-Belanger perspective.
3. An alternative to accounting monitoring and control is legislative agenda control as proposed by Bendor et al. (1985) and Miller and Moe (1983). In the Niskanen framework (1971), a bureau exchanges a lump-sum budget for a promised amount of output. However, the legislature sets the agenda and can demand certain price-quantity relationships. This agenda lets the authority of the legislature dominate and, potentially, overcome the information monopoly of the bureaucracy. The Bendor et al. (1985) model is based on the U.S. federal government and it is not clear that it will work effectively at the local level.
4. Most municipal debt in the U.S. is exempt from income tax and pays a lower interest rate than commercial counterparts. Most long-term local government debt in the U.K. is borrowed from the Public Works Loan Board, a government agency that borrows from the financial markets and then lends to local governments. About 10% of U.K. local debt comes from bank borrowing.
5. Most geographic areas in the U.K. have two local governments ("two-tier"), one at the city or town level and one at the county level. This is similar to U.S. governments. A unitary district is a geographic area in which there is only one local government ("single-tier"), essentially abolishing one level of government.
6. All private accounting firms in the U.K. sample are Big Six. The AUDITOR dummy variable attempts to measure the impact of differences (implicitly related to audit quality) between private firms and District Auditors.
7. U.K. policy experts contacted expected this finding, because of a policy shift by conservative governments toward the private sector, especially for seniors. Also, the National Health Service pays for medical costs, not the local districts.
8. Only a small percentage of revenue is under local control in the form of a local property tax. The central government has the legal authority to cap the size of local revenues. The same regression model used to distribute grants, determines the "standard spending assessment" for each authority. If the local district revenue is greater than the standard assessment, the central government can issue a cap and limit local spending.
9. Other governments may pick up some of the slack. School district spending has increased, as has public medical payments. However, these are associated with Medicare

and Medicaid in the U.S. (federal and state programs) and the National Health Service in Britain.

10. Also, the versions of the GM model are not identical. For example, the auditor variables test different things. The Big Six/non-Big Six variable is a common "brand name" or quality measure. However, British local audits are conducted either by District Auditors or Big Six firms. This dummy tests for differences in these two groups.

REFERENCES

Anthony, R., & Herzlinger, R. (1975). *Management control in nonprofit organizations*. Homewood: Richard D. Irwin.

Ashworth, J., & Gemmell, N. (1996). Local governments' monopoly power and local tax reform: Theory and evidence from the U.K. *Public Choice*, *89*(December), 393–417.

Bendor, J., Taylor, S., & Van Gaelen, R. (1985). Bureaucratic expertise vs. legislation authority: A model of deception and monitoring in budgeting. *American Political Science Review* (December), 1041–1060.

Breton, A., & Wintrobe, R. (1975). The equilibrium size of a budget-maximizing bureau: A note on Niskanen's theory of bureaucracy. *Journal of Political Economy*, *83*, 195–207.

Browning, E., & Browning, J. (1994). *Public finance and the price system* (4th ed.). New York: MacMillan.

De Angelo, L. (1981). Auditor size and auditor quality. *Journal of Accounting and Economics* (December), 183–199.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy* (April), 135–150.

Fawson, C., & Giroux, G. (1988). An empirical extension of the municipal monopoly model to provision of community infrastructure. *Public Choice* (April), 79–83.

Giroux, G. (1989). Monopoly power and monitoring: A test using the gonzalez and mehay model. *Public Choice* (October), 73–78.

Giroux, G., & Jones, R. (1998). Accounting and auditing for local governments in the U.S. and the U.K. Working Paper. Texas A&M University.

Giroux, G., & Shields, D. (1993). Accounting controls and bureaucratic strategies in municipal government. *Journal of Accounting and Public Policy* (Fall), 239–262.

Gonzalez, R., & Mehay, S. (1985). Bureaucracy and the divisibility of local output. *Public Choice*, *45*(1), 89–101.

Jones, R., & Pendlebury, M. (1991). The published accounts of local authorities, revisited. *Financial Accountability and Management* (Spring), 15–33.

Migue, J., & Belanger, G. (1974). Toward a general theory of managerial discretion. *Public Choice*, *17*, 24–43.

Miller, G., & Moe, T. (1983). Bureaucrats, legislators, and the size of government. *The American Political Science Review* (June), 297–322.

Mueller, D. (1976). Public choice: A survey. *Journal of Economic Literature* (June), 395–433.

Niskanen, W. (1971). *Bureaucracy and representative government*. Chicago: Abline-Atherton.

Niskanen, W. (1975). Bureaucrats and politicians. *Journal of Law and Economics*, *18*(October), 617–643.

Pendlebury, M., & Jones, R. (1985). Governmental budgeting as ex ante financial accounting – The U.K. case. *Journal of Accounting and Public Policy* (Winter), 301–316.

Toma, M., & Toma, E. (1976). Bureaucratic responses to tax limitation amendments. *Public Choice* (Spring), 45–61.

Tullock, G. (1965). *The politics of bureaucracy*. Washington: Public Affairs Press.

Wagner, R. (1976). Revenue structure, fiscal illusion and budgetary choice. *Public Choice, 25,* 45–61.

Weingast, B. (1984). The congressional-bureaucratic systemic principal agent perspective (with applications to the SEC). *Public Choice*, *44*(1), 147–191.

A LITERATURE REVIEW AND ACCOUNTING RESEARCH AGENDA FOR EDUCATIONAL PERFORMANCE INDICATORS

Marc A. Rubin

ABSTRACT

The development of educational performance and accountability measures and reports is a response to the demand by taxpayers, elected officials, and parents for information regarding the return on large amounts of resources invested in schools. Although a great deal of academic and professional effort has and continues to be expended on this topic, the accounting discipline has not been as active participant as other disciplines. Accountants are trained as information specialists and have the skills that can be useful in the development and improvement of educational performance and accountability reports. This paper reviews significant academic and professional efforts in developing educational performance information and offers an agenda for improving the production and use of educational accountability information.

1. INTRODUCTION

Traditionally, accountants and accounting researchers focus on the production and use of financial information for determining resource allocations. Public sector

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 239–261

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11012-4

accounting is primarily concerned with the development of budgets and provision of annual financial reports. These traditional reports are coming under scrutiny by decision makers who are questioning their usefulness in evaluating whether a government is effectively and efficiently delivering services. Among the tools that are used to overcome some of the perceived inadequacies of traditional reports are performance indicators.

Over a decade ago the Government Accounting Standards Board (GASB) demonstrated its interest in both financial and non-financial performance measures of government services by undertaking a project on service efforts and accomplishments (SEA). The GASB (1987) emphasizes the importance of SEA by stating in Concept Statement 1 that,

> Financial reporting should provide information to assist users in assessing the service efforts, costs, and accomplishments of the government entity. This information, when combined with information from other sources, helps users assess the economy, efficiency, and effectiveness of government and may help form a basis for voting or funding decisions (para. 77c).

The GASB (1994) specifically addresses the objectives of service efforts and accomplishments reporting in Concept Statement 2 as follows:

> The objective of SEA reporting is to provide more complete information about a governmental entity's performance than can be provided by the operating statement, balance sheet, and budgetary comparison statements and schedules to assist users in assessing the economy, efficiency, and effectiveness of services provided (para. 55).

Recently, the GASB demonstrated continued interest in performance measures with a web site (http://www.rutgers.edu/Accounting/raw/seagov/pmg/index.html) devoted to providing examples of governmental performance indicators as well as links to documents discussing a variety of issues related to performance indicators.

Although the GASB interest in performance indicators is evident, there is not a significant amount of accounting research addressing the production and use of performance indicators. The development and reporting of government performance indicators is becoming commonplace, yet many questions remain regarding the effect of performance information on the efficient and effective allocation of financial resources within and between government organizations. My purpose is to provide a literature review and discussion of accounting research issues related to accountability and performance indicators for one type of governmental entity, primary and secondary schools.

Education is a significant government service and appropriate for the study of performance indicators for a number of reasons. Schools are an important government-provided service based upon their societal purpose and consumption of large amounts of public resources. Thus, the interest in school productivity and accountability has and continues to be a salient topic. Elmore of Harvard's graduate school of education states that,

> Accountability for student performance is one of the two or three – if not the most – prominent issues in policy at the state and local levels right now (Olson, 1999a, p. 8).

Ladd supports this observation stating,

> Given that resources are not likely to increase much in the future, and there is a need for a better-educated work force, most people agree that schools need to become more productive in the future (1996a, p. 3).

In addition to their importance, educational services are typically provided by a separate single-purpose government entity. The widespread offering of educational services and the ability to easily determine the entity responsible for providing the services makes this service well suited for the study of government performance indicators.

In addition to accountants, other professional groups and disciplines are interested in educational performance indicators and accountability. A considerable portion of the research devoted to educational performance indicators is found in the economics of education and educational finance literatures. The research done in these areas contributes important insights that accounting researchers need to consider when analyzing educational accountability and performance information. The next section reviews literature relevant to educational accountability research area. Subsequent to the literature review, Section 3 describes a variety of proposals for educational performance indices and reports. Section 4 reviews the current practice of providing educational performance reports. Section 5 suggests a research agenda for the study of educational performance indicators from an accounting perspective. Concluding comments are offered in Section 6.

2. LITERATURE ON EDUCATIONAL PERFORMANCE REPORTING

Although spending in the United States on primary and secondary education is approximately $7000 per student totaling over $300 billion per year, American adults recently expressed the opinion that the quality of education in the U.S. remains a significant concern (National Public Radio/Kaiser/Kennedy School of Education, date unknown). Recently many public school districts are developing reports for stakeholders concerning the quality of educational services they provide. The inherent difficulties with measuring educational performance make it difficult for schools to develop comparable reports. Many states now require schools to provide a “report card” for use by constituents, but there is a lack of consensus on the characteristics of schools to measure, measurement techniques, and report formats (Olson, 1999b). For example, school efficiency, which ultimately may be the critical factor in raising and allocating educational resources, is

often not addressed in any significant manner in these school report cards. In order to assess the current value of educational performance reporting and to undertake necessary changes, we need to understand the users and uses of educational performance reports.

Most of the research on educational performance indicators is based upon economics and educational finance and administration disciplines. Accountants need to be aware and draw from the insights provided by these other disciplines if they are to make valued contributions advancing educational accountability and performance reporting. Information issues of interests to accountants are directly related to the underlying economic resource issues addressed in the economics and educational finance areas. The economics of education literature is primarily focused on determining the relationship between the resources provided to schools and the attainment of educational goals. That is, economists are interested in ascertaining the production function that relates education inputs to outputs. The development of an agreed upon education production function would allow decision-makers to make efficient resource allocations. Educational finance experts are also interested in the use of information for holding school officials accountable for resource decisions. In this section I review both the relevant economics and educational finance literatures since they are fertile areas for understanding the development of education performance information.

2.1. School Accountability and Performance Measures

Many of the issues that underlie educational performance reporting have been identified and discussed for a considerable period of time. For example, in one of the earlier papers focusing on educational performance, Levin (1974) proposes a structure for developing an educational accountability process. Levin suggests that there are four distinct underlying dimensions of educational accountability: as performance reporting; as a technical process; as a political process; as an institutional process. Levin believes that two problems surround using performance reports for accountability.

> One of the salient issues that arises is accountability to whom and for what. The performance-reporting approach avoids this dilemma by assuming that existing political processes will produce the appropriate goals for all parties, a highly questionable presumption.
>
> A second question raised by accountability as performance reporting is how the mere provision of information will provide "results." . . . There is the underlying hope that performance-reporting systems will lead to favorable changes and make the schools more responsive, but no mechanism for doing that is provided by the information itself (1974, p. 365).

Levin proceeds to examine all four goals of accountability and attempts to integrate them into a single framework or system. His analysis examines broadly many of the problems in assessing the provision of educational services. He concludes:

> Perhaps the most important implication of this review is to suggest that a significant tightening of the accountability linkages in education is probably impossible without substantial changes in the governing processes and organizations of the educational sector. Yet educators are fond of talking about accountability as a technical problem which does not require any major restructuring of institutions. Such a viewpoint may place the educator at center stage in the accountability movement, but is unlikely to make much of a difference in the overall functioning of the schools or society (1974, p. 388).

Since Levin published this article over 25 years ago, certain aspects of providing educational services have changed. For example, some school districts offer vouchers that may have an impact on the accountability process. Yet, many of Levin's insights are still applicable. More recently Wiggins (1993) shares many of Levin's concerns and believes that changes are needed to institute an accountability system that improves the provision of educational services. Wiggins concludes that,

> If faculty sets clear public targets, if parents have a clear and present voice, and if students, former students and institutional customers have a voice, then we will realize accountability that has the power to improve our schools (1993, p. 22).

Many of the concerns expressed by Levin were also reiterated in a series of articles based on a Brookings Institution conference (Ladd, 1996a). Ladd believes that educational reform requires attention be paid to goals and outputs rather than the traditional focus on inputs such as money allocated to schools. As did Levin, Ladd suggests that articulating goals is alone insufficient for accountability to impact learning. She states,

> The challenge within the educational system is to find the specific strategies and policy levers that will change enough parts of the complex educational system to increase student learning (1996b, p. 5).

Presenters at the Brookings conference offered two types of educational reform, increasing the accountability within existing administrative structures, and changing to entirely new structures and systems for delivering educational services. In either scenario, Ladd (1996b) suggests that reliable information is a necessary component of the accountability process.

As part of the Brookings conference, Elmore, Abelman and Fuhrman (1996) prepared analyses of specific educational accountability systems in two states. Based upon their study two of the five challenges identified with instituting educational accountability systems require improved or additional accountability

information. They observe one of the challenges of educational accountability is that systems need to be understandable and defensible to constituents. Information is at the core of this challenge. A second challenge is the need for stable political environments in order for educational accountability systems to succeed. A critical component of creating a stable political environment is the role the media plays in providing appropriate information to constituents.

Ferris (1992) assesses school accountability from a conceptual focus by using the principal-agent perspective. He specifically examines the potential costs and benefits of decentralizing curriculum, staffing and budget decisions from the district to the school level. Ferris acknowledges the difficulty of assessing school performance and placing responsibility on the parties contributing to performance outcomes. He summarizes the challenge of information as follows:

> Accounting and auditing schemes to monitor the expenditure of funds are fairly easy to devise. It is harder to develop schemes to ensure that the funds are being used to increase performance. But it is not clear the district has any advantage over the school site as to deciding on the best allocation (1992, p. 338).

Porter (1991) develops a framework for providing school indicators. He believes that indicators can serve one of three purposes; descriptive, monitoring school reform, and explanatory information when school output goals are not reached. Porter provides a model of the educational system that can be used as the basis for developing appropriate indicators. Based upon his model, Porter discusses many of the problems surrounding the development and use of educational indicators, including measurement and sampling issues. He then provides three criteria for selecting indicators: (1) relevance; (2) reliability and validity; and (3) benefits vs. cost. Porter suggests that because measurement issues remain problematic, indicators should not be used for accountability purposes. He then suggests a plan for the development and implementation of school process indicators.

2.2. Education Production Functions

In addition to the variety of conceptual articles that address school performance measures, there exists a literature on identifying the attributes of the educational process that lead to successful student achievement. The literature on educational production functions has the potential for guiding the selection of performance measures that most relate to student learning, which is the primary concern of school stakeholders. A significant portion of the educational production function literature is reviewed and synthesized by Hanushek (1979, 1986, 1991, 1997) and Monk (1992).

According to Hanushek (1991), most of the educational production function studies use achievement test scores, attendance rates, dropout rates, college matriculation and post-school earnings as measures of school performance. Educational inputs used in production function studies include a variety of socio-demographic characteristics, teacher characteristics, aspects of the school organization and community factors. Based upon his review of the education production function literature, Hanushek concludes that,

> Student achievement at a point in time is related to the primary inputs: family influences, peers, and schools. The educational process is also cumulative, so that both historical and contemporaneous inputs influence current performance (1997, p. 141).

He also states that, "There is no strong or consistent relationship between school resources and student performance" (1997, p. 148). He further clarifies this conclusion as follows,

> The concern from a policy viewpoint is that nobody can describe when resources will be used effectively and when they will not. In the absence of such a description, providing these general resources to a school implies that sometimes resources might be used effectively, other times they may be applied in ways that are actually damaging, and most of the time no measurable student outcome gains should be expected (1997, pp. 148–149).

Upon reviewing the research in educational productivity, Monk arrives at a conclusion similar to Hanushek's conclusion. He states that,

> In particular, there is an optimism about the prospects for improving educational productivity through the use of state fiscal policy that is remarkable given the disappointing results of the most recent policy-relevant research on the subject, namely the attempts to estimate education production functions (1992, p. 307).

Monk expresses his skepticism that school productivity can be enhanced without an improved understanding of the relationship between educational inputs and outputs. He suggests that incentive policies to improve schools can only be effective if the educational production function is articulated. Monk believes that future research on educational production functions will likely be more fruitful if classroom-oriented analyses are conducted instead of the more aggregated approaches assessing school, school district or aggregated state data.

The educational accountability and performance research literature provide guidance toward selecting the factors that have the potential to impact student success such as those suggested by Hanushek (1997). The research literature also clarifies the problems that exist in providing educational performance information, including the identification of the precise nature of the characteristics that affect student achievement, the current and cumulative effect of variables on students, the lack of input, output and outcome measurement precision, and the need to further determine how decision makers utilize performance information. In order

to develop and improve educational performance reports accountants need to consider including information on factors that relate to student success as well as being aware of measurement and communication problems. Further research on educational performance reporting needs to focus on resolving these issues. The aforementioned measurement problems and the lack of understanding of the relationship between educational inputs and outputs have not deterred proposed models of educational performance indicators or actual reporting of school performance. The next section reviews a few of the proposed models of school reporting and the subsequent section describes the current state of providing school performance information. Many of the proposed measurement models and indices as well as available accountability reports have common characteristics, although considerable variability remains.

3. PROPOSED EDUCATIONAL PRODUCTIVITY INDICATORS AND REPORTS

Boyer (1991) provides a framework for assessing school success by using a national education index. The intent of the index is to provide stakeholders in the educational system such as taxpayers, government officials, and parents of school-age children, additional information on the current status of the nation's schools. The index contains six components; student achievement, conditions of teaching, school climate, school finance, accountability and intervention, and building partnerships. Boyer's suggestions for measurable dimensions of each of these components are described in Table 1. Boyer believes that this index needs to be interpreted in context of the specific goals of the education providers. The index would be calculated by each state. Boyer does not indicate whether the index would be some form of aggregate information, but he does suggest that the implementation of an index should be left to a newly created national council.

Guthrie (1994) also addresses the issue of a composite indicator of school performance. Although Guthrie believes that indicators can be misused because they oversimplify the ability to accurately measure school performance, they are in demand by the media and public. Guthrie suggests four components of a school indicator; student performance, public support for education, conditions of children and educational service quality. The measurable dimensions of these components are summarized in Table 2. Guthrie suggests a method to aggregate the individual measures but describes unresolved issues that need to be overcome before the index is put into practice.

Education Week (Olson, 1999c) collaborated with a private group, A-Plus Communications, to study school report cards as part of their special issue on

Table 1. A Summary of Educational Performance Indicators to be Included in a National Education Index as Suggested by Boyer (1991).

Component	Measurable Dimensions
1. Student achievement	• Percent of students successfully finishing prescribed course of study • English proficiency • Math proficiency • General knowledge • Priority given to writing • Creativity and independent thinking skills
2. Condition of teaching	• Plan for attracting outstanding students to teaching • Special strategies to attract math and science teachers • Mandated in service programs • Starting and maximum salaries with regional comparisons
3. School climate	• Quality of buildings, classrooms and labs • Class size
4. School finance	• Per-pupil expenditures with cost-of-living considerations • Equity considerations in funding
5. Accountability and intervention	• Autonomy • Clearly-defined goals • Annual assessment and reporting procedures • Local schools held accountable for outcomes not procedures • Clearly defined methods of intervention if failure
6. Building partnerships	• Program for parent involvement in child's education • Promotion of support from business community

school accountability. Based upon the results from a series of focus groups and surveys, a prototype school performance report was developed. The report contains measures of student performance, ratios relating resources to the number of students, descriptions of student demographics and staff experience, data on spending and how classroom time is allocated, and information on school environment such as safety and parent involvement. An aggregate measure of performance is not proposed as part of the report. Details on the measures suggested in the prototype report are found in Table 3.

The Government Accounting Standards Board study on the reporting of service efforts and accomplishments of elementary and secondary education (Hatry et al., 1989) is likely the most significant contribution of the accounting profession to the development of educational performance indicators. As part of the study, the authors suggest information that would help interested parties evaluate the

Table 2. Summary of the Components of the Educational Composite Indicator as Suggested by Guthrie (1994).

Component	Measurable Dimension
1. Student performance	• Unstated
2. Public support for education	• Public opinion regarding school performance • Expenditures for schooling as a percent of personal income • School board election turnout • Mean teachers salaries as a percent of national mean personal income • College freshman's views of teaching as an occupation
3. Conditions of children	• Children's health • Mental health • Criminality • Poverty • Family stability • Personal attitudes and aspiration levels
4. Educational service quality	• School building modernity and upkeep • Availability of advanced placement courses • Measures of teacher quality • Length of school day and year • Availability of head start programs • Library or information resources • Teacher-pupil ratios

provision of educational services. They suggest measures of inputs, outputs, outcomes, and efficiency along with explanatory data that would provide users with the necessary information to evaluate performance of schools. Table 4 summarizes the information that is suggested for reporting. The GASB report suggests a wide range of measures, including indicators that have a financial component and others that do not include financial information. The GASB report also distinguishes between controllable and non-controllable factors. This dichotomy, which is often overlooked in the educational performance indicator literature, may ultimately prove to be critical in the appropriate use of performance indicators in an accountability system. The GASB report provides examples of selected state, school district and school reports containing performance information.

The GASB followed up on its series of research reports on service efforts and accomplishments with Concept Statement Number 2 (GASB, 1994). The Concept Statement provides a general framework for developing government performance information and encourages governments to experiment with

Table 3. The Components and Measures of the Prototype Report as Suggested by the Education Week Study (1999).

Component	Measure
1. Demographics	• Number of students • Number of teachers • Number of administrators • Number of classroom support staff • Number of school support staff • Percentage of students who attended preschool • Percentage of students with a home language other that English • Percentage of students qualifying for free or discounted lunch
2. Student performance	• Percentage of students who met goals on state mastery test • Attendance rate • Promotion rate
3. How money is spent	• Student to teacher ratio • Student to counselor ratio • Students to computer ratio • Average teacher experience • Percent of teachers with graduate degree • Percent of teachers trained as mentors • Number of certified teachers • Number of trainee teachers • Per-pupil spending on teachers and classroom materials • Per-pupil spending on counselors • Per-pupil spending on teacher training • Per-pupil spending on utilities and maintenance • Per-pupil spending on administration
4. School environment	• Number of suspensions per 100 students this year • Number of violent incidents per 100 students this year
5. How time is spent	• Number of hours of instruction per year in each area of instruction • Percentage of students in bilingual programs • Percentage of students in gifted programs • Percentage of students in special education

producing and using performance indicators and reports. GASB recognizes that before specific accounting and reporting standards can be issued regarding performance indicators the following issues still need to be explored,

(a) assess the types of SEA information that are useful,
(b) develop valid, acceptable measures of SEA,
(c) gather information required for SEA measures, and

Table 4. Summary of Information Suggested to be Included in School SEA Reports by the GASB Study (1989).

Component	Measure
1. Inputs	• Expenditures • Total number of personnel
2. Outputs	• Number of student-days • Number of students promoted • Carnegie units as percentage requirement • Absenteeism rate • Dropout rate and/or retention
3. Outcomes	• Types of tests • Test scores by major area • Measure of gain on achievement test • Measure of self-esteem • Measure of physical fitness • Measure of post-grad employment/education • Self-assessment by students of skills • Parent assessment of student skills
4. Efficiency (input to output or outcome)	• Cost per student (student, student-day) • Cost per outcome (cost per achieve) • Cost/program • Cost/school
5. Explanatory data	Controllable • Class size or pupil/teacher ratio Non-controllable • Attendance • Measure of minority students • Measure of student on reduced lunch • Measure of need for remedial programs • Student mobility rate measure • English as a second language • Student enrollment

(d) develop methods to report, explain, and verify information about SEA (GASB, 1994, para. 6).

More so than other literature on performance indicators, the GASB particularly encourages governments to supply performance information on outcomes, outputs and efficiency. The GASB has the future goal of issuing standards regarding the production and reporting of performance information that will be useful to the public for accountability and decision-making.

Although Boyer (1991), Guthrie (1994), *Education Week* (Olson, 1999a, b, c) and the GASB (Hatry et al., 1989) suggest common features to include in education performance reports, they differ somewhat in detail and focus. All of the suggested reports emphasize measures of student performance, school expenditures and class size. The GASB report suggests that performance information focus strongly on school outcomes, whereas Guthrie is more attentive to the condition of the individual child and Boyer to school processes. Both Guthrie and Boyer suggest an index that reflects the aggregation of the individual components, but Guthrie is somewhat more specific on how the aggregation should be accomplished. If we compare the suggested performance reports to the dimensions of educational productivity that the production function research finds important, the school environment dimension is given the most attention (particularly by Boyer). The other two dimensions, family influences and peer groups, receive considerable attention in the Guthrie model and to a lesser degree in the GASB report. The *Education Week* prototype is balanced between the various dimensions of educational productivity.

Users of school performance information desire comparable, understandable reports that provide information on the factors influencing student achievement and costs of providing services, but agreement on a common report is unlikely until previously described measurement and process issues are resolved. Models of reports, including aggregate measures of school performance must be developed cautiously. The next section reviews the current status of information being provided concerning school performance, demonstrating the variance in performance information communicated by schools.

4. CURRENT STATUS OF EDUCATIONAL PERFORMANCE REPORTS

Most of the recent focus on school performance reports is on the "school report card." Although the term "report card" is being used by most states to refer to a school or school district performance report, the types of information contained on these reports varies considerably between states. A recent issue of the publication *Education Week* (1999) contained a thorough state-by-state analysis of the information content of education report cards. For a school report card to be included in the *Education Week* study it must contain multiple indicators of performance assessment (1999, p. 87). The *Education Week* analysis categorizes the information contained in school report cards into the following groups of school characteristics; comparative student performance, academics and achievement, students, teachers, resources and school climate.

Table 5 summarizes the school report card analysis prepared by *Education Week*. As of 1998, 36 states require school report cards. A few additional states subsequent to the *Education Week* analysis have started to require school report cards. Table 6 contains a list of web sites for state school performance reports. All report cards include test scores, and either graduation and dropout rates or both. In addition, most report cards include selected student characteristics, class size or pupil-teacher ratios and a measure of student attendance. A variety of other attributes are provided with less regularity. Most of the state report card attributes assessed in the *Education Week* analysis are suggested as important performance information in the Boyer (1991), Guthrie (1994), *Education Week* (Olson, 1999c) or GASB (1989) report models. Current state report cards have a tendency to provide more information on school environment than family and peer groups. Other forms of information, such as financial and efficiency, are found with even less regularity.

U.S. News (1999) also prepared a performance report for selected high schools. Their analysis assesses four output or outcome measures; state test scores (math and English), percentage of students taking the SAT or ACT, advanced placement test taking (the average number of AP tests taken by seniors), and persistence rate (percentage of students in grades nine through 12 who complete the school year). Unlike most other performance reports, *U.S. News* develops a model of expected output/outcome results. They associate socioeconomic characteristics of a school's student population with the outcome measures. Outcome results based upon the model are compared with actual results for each school. Schools are then rated based upon this comparison, with the highest graded schools being those schools whose actual results most exceeded the expected results. *U.S. News* applies their model to a sample of urban high schools. The authors conclude that the following six school characteristics correlate with excellent performance: a challenging core curriculum, high standards and high expectations, highly qualified and well-trained teachers, family-school partnerships, mentors to motivate students, and high attendance. These characteristics overlap with the significant characteristics suggested in the education productivity research (family influences, peer groups and school environment). Many of these characteristics are reported in state report cards. *U.S. News* does not include any financial or efficiency measures in their analysis.

Currently a variety of school performance reports are available and additional methods of reporting school performance are being developed. For example, both PricewaterhouseCoopers (2000) and Standards & Poor's (2001) have recently announced initiatives in school accountability reporting. As mentioned, reasons for the variability in reporting is likely due to the lack of consensus on the critical characteristics affecting student performance and the problems in measurement. Also, the audience for school performance reports is not well specified and the process for using performance reports in decision-making is unclear. The next

Table 5. Summary of Results of the Education Week Study of 36 State School Report Cards (1999).

Dimension	Attribute	Number (%) of States That Include Attribute in Report Card
Publishing and Disseminating	Does the state have an annual report card on each of its schools?	36[a] (100%)
	Does the state make all the school report cards available on its web site?	26 (72)
	Does the state require the school report cards to be sent home?	13 (26)
Comparing student performance	National average	17 (47)
	State average	25 (69)
	District average	20 (56)
	Previous year	25 (69)
	Scores in similar schools or districts	9 (25)
	Scores predicted by student demographics	3 (8)
	Scores in top-performing schools	1 (3)
Academics and achievement	Test scores	36 (100)
	Graduation rate	24 (67)
	Dropout rate	33 (92)
	AP courses or tests	14 (39)
	Course-taking	11 (31)
	SAT/ACT data	20 (56)
	Post secondary plans/experiences	15 (42)
Students	Student characteristics	23 (64)
	Student mobility	12 (33)
Teachers, resources and school climate	Teacher qualifications	16 (44)
	Salaries or other financial data	17 (47)
	Safety of discipline	17 (47)
	Class size/pupil-teacher ratio	20 (56)
	Student attendance	30 (83)
	Parent involvement	11 (31)
	Satisfaction/opinion data	5 (14)
Other information about school	Accountability rating	9 (25)
	Description of programs/philosophy	15 (42)

[a] Five additional states will have reports starting in 2000 or 2001. Six other states provide individual school test score results.

Table 6. Individual State Web Sites Locations for School Performance Reports.

Alabama: www.alsde.edu
Alaska: www.eed.state.ak.us/stats/
Arizona: www2.ade.state.az.us/srcs/
Arkansas: N/A
California: http://www.cde.ca.gov/psaa/
Connecticut: N/A
Colorado: http://www.cde.state.co.us/index_assess.htm
Delaware: http://issm.doe.state.de.us/profiles/
District of Columbia: N/A
Florida: N/A
Georgia: www.doe.k12.ga.us, http://168.31.216.190/
Hawaii: N/A
Idaho: N/A
Illinois: www.isbe.state.il.us/research/reports.htm#Report%20Card
Indiana: N/A
Iowa: www.state.ia.us/educate/publications/coe.html
Kansas: http://www.ksbe.state.ks.us/reportcard.html
Kentucky: www.kde.state.ky.us/comm/commrel/school_report_card/
Louisiana: www.lcet.doe.state.la.us/doe/omf/sps9899/spsframe.asp
Maine: http://janus.state.me.us/education/lres/lres.htm
Maryland: http://www.msde.state.md.us/MSPReportCard/default.htm
Massachusetts: www.doe.mass.edu/mcas/99mcas/toc.html
Michigan: www.mde.state.mi.us/reports/msr/
Minnesota: http://cfl.state.mn.us/PUB&RES.htm
Mississippi: www.mde.k12.ms.us/account/report/mrc.htm
Missouri: www.dese.state.mo.us/reportsummary/
Montana: www.metnet.state.mt.us/Montana%20Education/OPI/Measurement&Acctability/
Education%20Profile/HTM/index.shtml
Nebraska: N/A
Nevada: www.nsn.k12.nv.us/nvdoe/
New Hampshire: www.state.nh.us/doe/Reports%20and%20Statistics/reports.htm
New Jersey: www.state.nj.us/njded/stass/index.html or www.state.nj.us/njded/reportcard/index.html
New Mexico: N/A
New York: www.emsc.nysed.gov/repcrd399
North Carolina: www.dpi.state.nc.us/accountability/reporting/index.html#Report
See also: http://www.smartschools.org/
North Dakota: www.dpi.state.nd.us/dpi/reports/publicat.htm
Ohio: www.ode.state.oh.us/rc_download.htm
Oklahoma: http://sde.state.ok.us/pro/stutest/drc.html
Oregon: www.ode.state.or.us/ReportCard/
Pennsylvania: www.paprofiles.org/
Rhode Island: N/A
South Carolina: N/A
South Dakota: www.state.sd.us/deca/DATA/99digest/
Tennessee: www.state.tn.us/education/rptcrd99/index.html
Texas: www.tea.state.tx.us/perfreport/account/
Utah: www.usoe.k12.ut.us/pr/facts.htm
Vermont: http://crs.uvm.edu/schlrpt/
Virginia: www.pen.k12.va.us/html/reportcard.shtml
Washington: www.k12.wa.us/assessment/default.asp
West Virginia: wvde.state.wv.us/data/report_cards/
Wisconsin: N/A
Wyoming: N/A

section suggests a research agenda that may assist researchers in addressing unresolved issues in school performance reporting.

5. ACCOUNTING RESEARCH AGENDA ON EDUCATIONAL PERFORMANCE REPORTING

Ultimately information value depends on whether it has an effect on the decision-making behavior of individuals. Educational performance information will only continue to be provided if perceived benefits (ability to improve decisions) exceed costs. There is increasing interest and resources devoted to educational accountability and performance-reporting systems in the last decade. Yet, many issues need to be resolved if stakeholders are to make appropriate decisions with consequences that they desire. Although Porter (1991) develops a performance indicator system, he suggests that educational measures are not sufficiently precise to be useful for accountability. Porter's assumption regarding the lack of precision in educational performance measures may have merit, but the reporting of educational performance information will likely affect the ability of school officials to raise, budget, and spend resources whether or not that is the intent of the information providers. For example, taxpayers are often asked to decide on the merits of school tax levies and will therefore use any information available to them. In many cases the quality of the available information in making such decisions is overlooked. Therefore, the problems that exist with educational performance information need to be addressed promptly.

Monk (1997) addresses the problems of collecting and using school data. His analysis examines the information used for researching the origination, disposition and utilization of school resources. Monk suggests that education resource utilization information is currently very scarce. He states, "We know virtually nothing about the interface between the allocation of student time and effort on the one hand and the allocation of schooling resources on the other" (Monk, 1997, pp. 306–307). Monk also believes that decision makers underutilize the data currently collected and available. He states, " This is a particularly serious problem at the local district level since districts often lack the capabilities and/or interest in drawing upon existing data for decision-making purposes" (Monk, 1997, pp. 307–308). Monk believes that individual school officials are even less equipped for using education performance information for decision-making. Monk provides recommendations for future data collection including increasing attention to: differences in school and school district organizational structures, differences in the role played by regional service delivery mechanisms, differences

between states in the linkages between education and the delivery of other social services to children, differences in the cost-effectiveness of educational resources, and differences in the utilization of educational resources.

Fuhrman (1999) states that educational performance measures are already being used for accountability purposes in many states. She reviews the progress of educational accountability reports and suggests a variety of issues that need to be addressed in the implementation of the measurement systems. The issues that Fuhrman believes need to be addressed include the reliability and validity of measures, standards of performance, appropriate teacher incentives, complexity and fairness, stability and credibility of the accountability system, and the ability of the accountability system to ultimately impact the quality of student learning. A number of the issues suggested by Fuhrman are appropriate for accounting researchers to examine.

Accounting researchers can address many of the educational performance information issues described by Porter (1991), Monk (1997) and Fuhrman (1999). Academic accounting researchers addressing the following questions will help clarify the establishment of a research agenda that can enhance the production and use of educational accountability and performance information.

What type of education performance reports should be provided to constituents?
Who is to be held accountable based upon the information provided in the reports?
What types of decisions will be based upon the information provided in the reports?
What types of analysis will be performed on the information provided in the reports?

One of the first steps in establishing a research agenda is to specifically define the uses and the users of educational performance information. Most issuers of current educational performance reports do not clearly and directly articulate the purpose of the report. Though many educational performance reports such as school report cards are similar in nature, it is often not clear which constituents are targeted as the users of the report. Even if we assume that school district property owners and parents of school children are the primary groups for which the reports are written, we do not know whether the intended users have the requisite knowledge to analyze the report information. For performance information to have significant benefit, we need to determine whether school stakeholders likely have the familiarity and means to bear the necessary costs to become involved with the school taxing and budgeting processes. A research agenda on educational performance reporting needs to include studies that address these issues.

Individuals or groups being held responsible for educational performance are typically not clearly identified in the school reports. If school performance is "everyone's" responsibility, then whose actions and decisions should be evaluated to determine what needs to be changed if improvement is desired? Research needs to help in ascertaining the specific uses of performance reports and what likely benefits the reports can provide to users. The *Education Week* study on school report cards noted numerous instances where parents and educators had divergent ideas on the type of information that should be included in the report (Olson, 1999b). The lack of agreement between parents and educators may be due to the perceived differences in the use of the report cards. Whereas parents may be looking for information about the success of their children, educators may see the report card as a means for punishing them for results upon which they have only partial influence.

A variety of research techniques are potentially useful for providing evidence relating to the uses and users of educational performance reports. For example, interviews or surveys of state and local school officials responsible for report card programs can determine who the intended audience is for the report cards and how they expect the report cards to be used by the audience. Behavioral studies with actual report users could be conducted with educational performance information to determine the ability of the intended audience to process and analyze the information contained in school report cards and similar documents. In addition, the target audience could be surveyed and interviewed to ascertain perceived costs and benefits of the information in their jurisdiction. Empirical research techniques that capture actual behaviors and decisions could also be used, although this may depend on the length of time the reports have been produced and made available. Empirical evidence on the costs and benefits of educational reports depends on the availability of a variety of data. For example, data relating to the information collection and reporting costs, the target audience activity caused by the reaction to the provided information (such as votes, budget testimony, and calls to school board officials) and educational resource allocations over a period of time would be useful in hypothesis development and testing of the benefits of information such as school report cards.

What characteristics of education services do stakeholders need for making their decisions?

If we can determine the intended users and uses of educational performance information, then the types of information that are most salient for use by the target audience can be assessed. Much of the research related to educational productivity and performance reporting is applicable for accomplishing this task. Research may focus on the application of the economic-based research on educational production functions or assess the appropriateness of reporting models already in use by

states or other organizations (such as *U.S. News*). In addition to the economic-type studies, evidence on this issue can be collected from taxpayers, voters, parents of school-aged children, and school officials by using survey, behavioral, and other empirical methods. For example, Monk et al. (1998) used focus groups and surveys to collect evidence on the differences in perceptions between state officials and the public on the criteria for describing school district quality.

How do we measure the required characteristics of the educational services provided by government?

Once the relevant characteristics of the learning environment are identified, then research should focus on measurement alternatives most appropriate for the intended audience. Accounting research can help in identifying strengths and weaknesses of available alternative measurement techniques. This would include assessing the alternative measures for their relevance, reliability, validity and cost. The potential benefits and costs of aggregate measures should specifically be addressed. Many states include school "grades" on school report cards based on an aggregation of data measuring a variety of dimensions of education. Are these measures valid? Can these aggregate measures capture the "true" status of a school in a single measure? Are there significant potential problems in using aggregate measures?

Accounting research also can enhance the development of appropriate efficiency measures. Although the GASB report (Hatry et al., 1989) recommends efficiency measures for assessing schools, this is the area that has been most ignored by school performance reports. A possible reason that efficiency measures are often overlooked is the difficulty in developing suitable measures relating outputs to inputs. Yet, if decision-makers want to prioritize the allocation of available resources, they need information on what specific resources are the most productive and provide the "best" educational value. Empirical, analytical, or experimental research techniques can potentially improve our understanding of the utility of current efficiency measures and help to determine the benefits of alternative measurements that are currently not in use. Accounting researchers can extend the managerial accounting literature on cost drivers, activity-based costing and similar topics to determine "best" measures of inputs to outputs and outcomes.

How do we format the needed information to maximize its appropriate use?

The format for school report cards is discussed in the report on focus groups organized by *Education Week* (Olson, 1999b). Feedback suggests that users of reports want information that is sufficiently detailed for making well-informed decisions, but at the same time succinct. This is similar to other government

services. For example the GASB has published research reports on alternative reporting formats (Wilson, 1990) and popular reporting (Carpenter & Sharp, 1992). This issue is particularly salient today with the advent of technology and information dissemination over the Internet. Behavioral research techniques may prove to be the most beneficial in gathering evidence on this issue. As Internet reporting expands, data collection techniques may be able to use technology, such as specialized software programs, to provide information on the attributes of users accessing and analyzing web-based information.

Do the benefits of providing these reports exceed the costs of developing and disseminating the reports?

Research evidence on this issue, which is ultimately the question we want to address, is the aggregation of research conducted on all of the previous issues. In order to determine costs and benefits of educational performance reporting we need to know the users and uses of the information, the measurement and reporting alternatives, and the process used for analyzing the information and translating that to decisions. Accounting researchers as part of this effort should design studies that ascertain the factors that motivate school officials to provide performance information. If we assume that schools providing performance information determined that the benefits of providing this information exceed costs, we can then design empirical studies to discover the common factors that relate to the costs and benefits of providing school performance information. Costs and benefits of providing performance information can also be determined by examining the effects of providing such information on decisions such as tax levy and bond voting, budgeting decisions, school board elections, and location decisions of homeowners and parents of school-age children.

6. CONCLUDING COMMENTS

Education receives a significant amount of public resources and there is a growing demand for accountability for those resources. While educational performance measurement and reporting is the focus of a considerable amount of economics and educational finance research, accountants have provided limited contributions. Many issues remain that require further evidence in order for us to understand the role this information has in the allocation and use of financial resources to operate educational institutions. Accounting research has traditionally been concerned with the relationship between information and financial resource allocations. Research on the development and impact of educational performance information is a logical extension of prior accounting research. If educational performance

measures become a significant, if not primary, source of information used for resource allocation decisions it then becomes necessary for accounting to directly address this form of information reporting in order to remain a viable and relevant discipline for producing and disseminating school information. My purpose is to provide both background and a structure for conducting accounting research in educational performance information.

REFERENCES

Boyer, E. (1991). The national education index. *Vital Speeches*, *57*(January 1), 189–192.

Carpenter, F., & Sharp, F. C. (1992). *Popular Reporting: Local Government Financial Reports to the Citizenry*. Norwalk, CT: Governmental Accounting Standards Board.

Education Week. (1999). *Quality Counts '99 Rewarding Results, Punishing Failure*, *18* (January 11).

Elmore, R. F., Abelmann, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. In: H. F. Ladd (Ed.), *Holding Schools Accountable Performance-Based Reform in Education* (pp. 65–98). Washington, DC: Brookings Institution.

Ferris, J. M. (1992). School-based decision making: A principal-agent perspective. *Educational Evaluation and Policy Analysis*, *14*(Winter), 333–346.

Fuhrman, S. H. (1999). The new accountability. *CPRE Policy Briefs*, *RB-27* (January), 1–11.

Government Accounting Standards Board. (1987). *Concept statement number 1: Objectives of financial reporting*. Norwalk, CT.

Government Accounting Standards Board. (1994). *Concept statement number 2: Service efforts and accomplishments reporting*. Norwalk, CT.

Guthrie, J. W. (1994). Do America's schools need a 'Dow Jones index'? *The Clearing House*, *68*(Nov/Dec), 98–103.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, *14*, 351–388.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, *24*(September), 1141–1177.

Hanushek, E. A. (1991). When school finance 'reform' may not be good policy. *Harvard Journal on Legislation*, *28*, 423–456.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, *19*(Summer), 141–164.

Hatry, H. P., Alexander, M., & Fountain, Jr., J. R. (1989). *Service efforts and accomplishments reporting: Its time has come elementary and secondary education*. Norwalk, CT: Government Accounting Standards Board.

Ladd, H. F. (Ed.). (1996a). *Holding schools accountable performance-based reform in education*. Washington, DC: The Brookings Institution.

Ladd, H. F. (1996b). Introduction. In: H. F. Ladd (Ed.), *Holding Schools Accountable Performance-Based Reform in Education* (pp. 1–19). Washington, DC: The Brookings Institution.

Levin, H. M. (1974). A conceptual framework for accountability in education. *The University of Chicago School Review*, *82*(May), 363–391.

Monk, D. H. (1992). Education productivity research: An update and assessment of its role in education finance reform. *Educational Evaluation and Policy Analysis*, *14*(Winter), 307–332.

Monk, D. H. (1997). Challenges surrounding the collection and use of data for the study of finance and productivity. *Journal of Education Finance, 22*(Winter), 303–316.

Monk, D. H., Nusser, J. L., & Roellke, C. R. (1998). Resource-based indicators for educational accountability and school district improvement. *Education and Urban Society*, *30*(August), 479–501.

National Public Radio/Kaiser/Kennedy School of Education Survey, available at www.npr.org/programs/specials/poll/education/education.results.html, date unknown.

Olson, L. (1999a). Shining a spotlight on results. *Education Week*, *18*(January 11), 8–10.

Olson, L. (1999b). Report cards for schools. *Education Week*, *18* (January 11), 27–28, 32–36.

Olson, L. (1999c). A closer look: What makes a good report card? *Education Week*, *18*(January 11), 29.

Porter, A. C. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis*, *13*(Spring), 13–29.

PricewaterhouseCoopers. (2000). *PwC and SchoolMatch announce alliance to provide management consulting and information services to the K-12 education market*. Press release at http://www.pwcglobal.com/extweb/ncpressrelease.nsf/DocID/4EBF76FEBE2A8153852568AA0066DE25 (March).

Standards & Poor's. (2001). *What is SES?* At http://www.standardpoor.com/ProductsAndServices/CreditMarketServices/SchoolEvaluationServices/School.html.

U.S. News Online. (1999). *Outstanding high schools: Lessons from six metro areas*. At http://www.usnews.com/usnews/issue/990118/18intr.htm (January 18).

Wiggins, G. (1993). Accountability, testing, and schools: Toward local responsibility and away from change by mandate. *Business Horizons*, (September-October), 13–23.

Wilson, E. (1990). *Financial reporting by state and local governments: A survey of preferences among alternative formats*. Norwalk, CT: Governmental Accounting Standards Board.

SCHOOL DISTRICTS USE OF THEIR WEB PAGES FOR FINANCING REFERENDA INFORMATION DISCLOSURE: AN EXPLORATORY ANALYSIS

Dennis M. Patten and Joel M. Strong

ABSTRACT

We investigate the extent to which 72 school districts in Illinois with financing referenda on the March 19, 2002 election ballot used their web pages for disseminating referenda information. Results indicate that while more than half (43 of the 72) of the districts did include at least some referenda disclosure on their web sites, the extensiveness of disclosure, based on content analysis, was somewhat limited. Contrary to expectations, we found no differences in the level of web page disclosure for districts with tax as opposed to bond referenda, or for urban and suburban as opposed to rural-based school districts. Larger districts did have higher levels of disclosure than smaller school districts.

Results of tests examining the relation between web page referenda disclosure and election outcome indicate that the value of disclosure varies across differing types of school districts and differing types of referenda. Greater web page disclosure was significantly associated with referenda success for

Research in Governmental and Nonprofit Accounting
Research in Governmental and Nonprofit Accounting, Volume 11, 263–282

ISSN: 0884-0741/doi:10.1016/S0884-0741(04)11013-6

districts with tax referenda but not for districts with bond referenda, for urban and suburban school districts but not for rural school districts, and for larger school districts but not for smaller school districts. These exploratory findings suggest that further investigation into the value of web page disclosure by school districts is warranted.

If the members of any community are going to govern themselves successfully they must have easy access to a reasonably accurate and reasonably complete account of the facts needed for intelligent decisions upon the issues facing them (Carl Cohen, *Democracy*, 1971, p. 56).

INTRODUCTION

Perhaps nowhere in direct democracy is the need for citizen access to accurate and complete information more important than in the case of school district financing referenda. Most states require school districts, which exist as local level governmental units, to obtain citizen approval for either bond issuances, tax levy changes or both (see Hamilton & Cohen, 1974, p. 3). Citizens are given the opportunity to either approve or vote against these financing options through public elections known as financing referenda. As summarized by Lentz (1999, p. 459) "School funding issues cut to the center of American society and democratic values, but raising money for schools is neither easy nor popular." Indeed, Lentz (1999, p. 459) reinforces this point by noting that during the 1980s Illinois voters rejected over half of the education fund tax referenda placed before them.[1]

Bowler and Donovan (1998), building on a rich stream of research on voting and voting behavior (e.g. Downs, 1957; Lupia, 1994; Magleby, 1984; Popkin, 1991), argue that information can play a key role in the passage of referenda. Our study examines the extent to which Illinois school districts are using their web pages as a vehicle for providing referenda information to potential voters. We also examine whether differences in the extent of this disclosure are associated with election outcomes.

We examined the web pages of 72 Illinois school districts with either an education fund tax referendum or a building bond referendum on the March 19, 2002 election ballot. Our review indicates that, contrary to our expectations, the likelihood of having web page disclosure and the extent of that disclosure did not vary across districts floating tax as opposed to bond referenda, nor across rural as opposed to urban or suburban based districts. As expected, larger school districts used web page disclosure more than smaller districts. However, our analysis finds that the relation between web page disclosure and referenda success varies across systematic factors. Higher levels of disclosure are associated with election

success for tax referenda but not bond referenda, for urban and suburban districts but not for rural districts, and for larger districts but not for smaller districts. These findings are all consistent with our hypotheses and suggest that the value of the Internet as a referenda communications medium for school districts is not uniform. We begin by developing the theoretical justification for our expectations.

BACKGROUND AND HYPOTHESIS DEVELOPMENT

Anthony Downs, in his seminal work *An Economic Theory of Democracy* (Downs, 1957), models party voting as an attempt to maximize one's utility given a world of uncertainty and limited information. As summarized by Downs (1957, pp. 46–47), "a man's evaluation of each party depends ultimately upon: (1) the information he has about its policies; and (2) the relation between those of its policies he knows about and his conception of the good society." However, because Downs assumes (p. 47) that every voter "has a fixed conception of the good life" and has already related it to his existing knowledge, he argues that "only new information can persuade him to change his mind" (p. 47).

Downs' model was developed relative to party election choice, but it has been adapted to the direct-democracy setting (of which school district financing referenda is a sub-set) by Magleby (1984), Bowler and Donovan (1998), and others. For our analysis, we assume Bowler and Donovan's (1998, p. 69) representation of direct-democracy voting "as a decision between an uncertain outcome [the world state if the referendum passes] and a more certain outcome (the status quo)." As summarized by Bowler and Donovan (p. 35), information is important in this decision context because "if voters approach a choice . . . with limited information, and they are uncertain about how the proposition might change the status quo, they can compare what they know about any *yes* vote with the potential certainty of preserving the status quo with a *no* vote." Thus, limited information is assumed to lead to a predisposition to vote *no*.[2] Based on Downs' argument that it is possible to change a voter's mind only by providing him with new information, Bowler and Donovan assert that a major focus in direct-democracy campaigns should be on better information provision.

A review of financing referenda resource literature suggests that school districts are aware of the importance of providing voters with information. Hamilton and Cohen (1974, p. 137), for example, assert that funding campaigns "are viewed by school forces as a gigantic communication offensive." Freeman (1990, pp. 26–27), citing Sissy Henry, the Director of Board Development in South Carolina and a workshop leader on passing financing referenda, stresses the most important thing to remember in developing a referenda campaign is to present voters "with

enough factual information so they can make a wise decision." Numerous recent referenda resource guides (e.g. Boer, 2001; Boschee & Holt, 1999; Graves, 1999) also recommend that those school districts wishing to pass referenda should concentrate on information dissemination. What is at least somewhat surprising, however, is that only one of these more recent guides (Boer, 2001) discusses the potential use of the Internet as a disclosure option.[3]

In spite of its limited life, the Internet is already being recognized in many circles as a powerful tool for information disclosure. To illustrate, both Ashbaugh et al. (1999) and Wildstrom (1997) suggest that the web can be used by corporations to communicate timely financial information to employees, shareholders, and others interested in the financial performance of the companies. Lally (1996, p. 1222) more emphatically argues that a "free market does not exist unless all parties can make their decisions based upon unlimited access to information" and that Internet technology allows for greater information dissemination. Jones et al. (2000, p. 71) reinforce the importance of this disclosure by noting that corporate stakeholders are increasingly using the web for gathering information. Within the governmental context, both Balutis (2001), addressing federal government use, and Hammel (2001), writing from a local governmental unit perspective, suggest there may be significant value to making information available to the citizenry over the Internet.

Clearly, the major advantage to school districts making disclosures on financing referenda available on the Internet is the potential for disseminating greater information to a broader audience. Because, as noted by Jones (2000, p. 71), there are virtually no space restrictions on the web, detailed information on specific referenda can be posted. For example, Boer (2001) suggests that justification for the referendum, cost estimates, projected operating budgets vis-a-vis past operating results, data on past funding requests, and examples of the actual ballot or ballot wording all might be considered for disclosure by the school district. The hyper-link capabilities of the web allow school districts to provide significant levels of detail in each of the areas, and thus allow users to access as much, or as little of the information as desired. The medium also allows the school district to respond to questions and/or criticisms on a timely basis. Although not explicitly recognizing these potential benefits, Boer (2001, pp. 21–22) does specifically identify using "Internet services to better inform school district residents of pertinent referendum information" as an essential component for a successful campaign.[4]

Unfortunately, little data are currently available on the extent to which the Internet is being used for information disclosure by school districts.[5] Recent surveys of governmental Internet use (e.g. Hammel, 2001; Moulder, 2001; Norris et al., 2001) examine city and county governments as opposed to school districts and tend to focus on the electronic delivery of community services and e-procurement rather than on information disclosure issues. Boer (2001) provides anecdotal evidence of school district use of web pages for referenda disclosure,

but provides only limited information. Accordingly, the first purpose of this study is to examine the extent to which school districts are currently using the Internet as a medium for providing information on pending financing referenda to aid citizens in the decision making process.

Beyond just compiling a descriptive analysis of the use of the web for financing referenda information disclosure, however, we also test whether greater levels of this disclosure are related to differences in voter outcomes.[6] A key premise underlying our analysis is that while providing more information via the school district web page increases the likelihood of successful passage, its value is not uniform. Procuring and evaluating information requires an investment of time and energy (Popkin, 1993, p.18). Thus, voters in Downs' rational choice model are not assumed to necessarily consume all of the information that may be available to them. Popkin (1993, p. 18) explains, for example, that "the collective nature of the vote means that there is low incentive for an individual to collect information solely in order to cast one vote among many." Instead, voters have an incentive to reduce their costs by relying on the use of what Popkin (1993) refers to as shortcuts, or information cues. And while the shortcuts most often discussed relative to party elections (party affiliation, past party performance) are not available in referenda-type choices, other types of information gathering shortcuts are assumed. Lupia (1994, p. 63) explains, for example, that "voters can acquire information about the preferences or opinions of friends, coworkers . . . or other groups, which they may then use" to form their own opinions on the issue. This suggests that where other types of shortcuts or information cues exist, the value of Internet information disclosure will be lower.

Based on the discussion presented above, we conjecture that the value of the Internet as a disclosure medium varies across certain systematic factors. The variation in these factors in turn influences both the level of disclosure of referenda information and its value in influencing voters. Each of these items is discussed below.

Referendum Type

First, we hypothesize that web page disclosure of information relative to financing referenda will vary across the type of referenda. For at least the recent past, bond referenda in Illinois have been substantially more successful at the ballot box than tax referenda.[7] Rado (2002, p. 8) reports, for example, that from 1997 to 2001 approximately 70% of Illinois school districts' bond referenda passed in contrast to only about 40% of the tax referenda. This may be due to the concrete evidence (overcrowded schools and/or deteriorating facilities) serving as an effective information cue for the need for building bonds. In contrast, such obvious

information cues tend not to exist relative to the needs driving tax referenda. Further, and in contrast to the more temporary impact of bond issuances, tax rate increases have no termination date and thus will continue to impact local property tax bills (Boer, 2001, p. 4). For these reasons, school district administrators may believe that greater justification for tax referenda is necessary to move voters away from the status quo. Concurrently, it seems likely that additional justifying information will have more impact where constituents are otherwise undecided, or even leaning against a proposal. Based on this, the value of Internet disclosure is predicted to be greater for justifying tax referenda than for bond referenda. Accordingly, we posit the following hypotheses (in alternative form):

$H1_a$. School districts floating Education Fund tax referenda will be more likely to include referenda disclosure on their web pages than school districts floating building bond referenda.

$H1_b$. School district web page disclosure of referenda information will be more extensive for districts with tax referenda than for those with bond referenda.

$H1_c$. There will be a greater correlation between school district web page referenda disclosure and referenda outcome for tax referenda than for bond referenda.

School District Classification

Among other breakdowns, the Illinois State Board of Education (ISBE) classifies each of Illinois' school districts according to its community-type base. Three classifications, urban, suburban, and rural are used.[8] We argue that information dissemination strategies are likely to be different for rural districts in comparison to urban and suburban based districts. Both Dewees (1999) and Hobbs (1995) assert that because rural school districts usually serve smaller, close-knit communities the school plays an important role as a community center and source of community pride, and thus voters in these areas are more likely to be aware of school district affairs. The need for referenda information disclosure may thus be lower in rural based districts. Further, as argued by Dewees (1999), closer community relationships also can make it easier to communicate with the local electorate. Both Boschee and Holt (1999, p. 23) and Boer (2001, p. 15) argue that personal contact with potential voters is a very effective means for promoting financing referenda, and Boer (p. 15) specifically encourages school board members in smaller communities to personally contact individual voters.

Based on these observations, it seems plausible that rural school districts, which are based in smaller communities, may see less value to using the Internet

for making referenda disclosures. Alternatively, metropolitan-based urban and suburban districts, which do not enjoy the benefits of the small town atmosphere to the same extent, may believe that the Internet is an efficient tool for disseminating their referenda message.[9] Further, and based on the arguments of Dewees (1999), voters from rural districts may be more likely to receive referenda information from friends and neighbors than their counterparts in urban and suburban settings, suggesting that Internet information may have more value in metropolitan-based districts. As such, we offer the following additional hypotheses:

H2$_a$. Urban and Suburban school districts will be more likely to include referenda disclosure on their web pages than rural school districts.

H2$_b$. Urban and suburban school district web page disclosure of referenda information will be more extensive than rural districts' disclosure.

H2$_c$. There will be a greater correlation between school district web page referenda disclosure and referenda outcome for urban and suburban school districts than for rural school districts.

School District Size

Previous studies of web page information disclosure document a significant relation between entity size and the extent of disclosure. This includes the disclosure of financial information (Ashbaugh et al., 1999), social responsibility information (Patten, 2002), and environmental information (Patten & Crampton, 2004). And although not examining disclosure issues, Hammel (2001) reports that larger non-metropolitan communities are more likely to have a local government web page than smaller communities, and that the use of the Internet is greater for the governmental units of the larger communities. Based on the results of these prior analyses, it seems likely that school district size may also be associated with the level of Internet disclosure of financing referenda information.[10] Further, and similar to the argument for differences in the value of web page disclosure across school district classification, smaller school districts may be in a position for more face-to-face contact with constituents and thus see less need for Internet dissemination of referenda information.[11] Finally, and again similar to the argument for rural school districts, we conjecture that disclosure will have more value for voters from larger rather than smaller districts. Our third set of hypotheses is thus stated as:

H3$_a$. Larger school districts will be more likely to include referenda disclosure on their web pages than smaller school districts.

$H3_b$. Web page disclosure of referenda information will be more extensive for larger school districts than for smaller districts.

$H3_c$. There will be a greater correlation between school district web page referenda disclosure and referenda outcome for larger school districts than for smaller school districts.

RESEARCH METHOD

Sample

The website of the Illinois Association of School Administrators (IASA) (www.iasaedu.org) provides the election results for the 100 Illinois school districts that had placed financing referenda on the March 19, 2002 ballot. By contacting the circuit clerk's office for each of the state's 102 counties in late January of 2002,[12] we had been able to identify all but one of those districts.[13] Because of our a priori expectations for differences in Internet disclosure due to referendum type, we chose to include in our sample only school districts with either an education fund tax levy or a building bond referendum on the ballot.[14] Accordingly, 12 districts were eliminated because they had multiple referenda on the ballot, two districts were eliminated because their referenda were for the operation and maintenance fund as opposed to the education fund, and two districts were eliminated because their requested bond issuances were not for building bonds. We were unable to either find, or access, web pages for 11 of the remaining districts. These school districts were also eliminated from our study resulting in a net sample of 72 school districts.[15] Thirty-nine of the districts had building bond referenda on the ballot and 33 districts were floating education fund tax levy referenda. Table 1 provides summary data on these school districts.

Web Page Referenda Disclosure Measures

We used content analysis to quantify the extent of referenda disclosure included on the sample school districts' web pages. Content analysis has been used in a number of disclosure-related studies (see, e.g. Patten, 2002; Wiseman, 1982) and involves reviewing the web pages for the presence or absence of various items of disclosure related to the referenda. We could find no studies that tested the relation between specific information availability and referenda success or failure. Accordingly, we relied on recent referenda resource guides (e.g. Boer, 2001; Boschee & Holt,

Table 1. Descriptive Statistics for a Sample of 72 Illinois School Districts with Financing Referenda on the March 19, 2002 Ballot.

Total illinois school districts with financing referenda on the March 19, 2002 ballot	100
School districts eliminated from sample (by reason)	
Not identified prior to election	1
Multiple referenda on ballot	12
Tax referenda not for education fund	2
Bond referenda not building bonds	2
Unable to find/access web page	11
Net sample	72
Sample school districts with building bond referenda	39
Sample school districts with education fund tax referenda	33
Sample school districts classified as urban or suburban[a]	52
Sample school districts classified as rural	20
Mean size of sample school districts (number of students)[b]	3,004
Median size of sample school districts (number of students)	1,547
Sample school district size ranges (number of students)	
Smallest	203
Largest	18,762

[a] Classification based on Illinois State Board of Education (ISBE) data available through the ISBE website (www.isbe.state.il.us).
[b] Based on number of students as identified in the 2000–2001 ISBE State Report Card File (available through the ISBE website).

1999; Graves, 1999) and informal interviews with school district administrators[16] to identify information areas that might be expected to influence the decisions of voters.

Separate 10 point content schemes were developed for analyzing the extent of disclosures related to tax referenda as opposed to bond referenda. However, most of the content areas examined were consistent across the scales. These included: (1) narrative justification for the referendum (coded zero to 3 depending on the extensiveness of the disclosure[17]); (2) the use of budget or actual financial data for justification of the referendum (coded zero for no use, one for the use of summary data only, and 2 for the use of detailed financial data); (3) disclosure of past referenda requests or tax increases (one point); (4) disclosure of the impact of the referendum on individual tax payments (one point); and (5) the provision of a sample ballot or ballot wording (one point).

Because of the differing nature of tax and bond requests, there were two differences in the respective content schemes. For tax referenda, points were awarded for the disclosure of comparative tax rate data (one point if either

multi-year comparisons or comparisons with other school districts were provided, two points if both types of comparisons were provided). For bond referenda, one point was awarded if project cost estimates were disclosed on the web page and one point was awarded for the provision of blueprints or architectural drawings.

Because Internet presentation is subject to constant change, and because the survey of Internet use for financing referenda information is somewhat exploratory, each of the sample school district web pages was reviewed on two separate occasions. The first review took place during the first two weeks of February, 2002, with the second review during the week preceding the election. For our analyses, we used the content scores from the latter review.[18] Appendix identifies the number of school districts making specific types of information disclosure across referendum type.

The statistical significance of differences in the extent of disclosure across classification schemes (tax versus bonds, urban/suburban versus rural, larger versus smaller) is measured using the non-parametric Mann-Whitney test. χ^2 tests are employed to measure for the significance of difference in election outcome percentages. Finally, both parametric (Pearson product-moment) and non-parametric (Spearman's rank-order) correlation techniques are used to test for the significance of relation between disclosure extensiveness and election outcome. Two separate measures of election outcome, a one/zero indicator variable (where one designates passage of the referendum) and the percentage of "yes" votes, are used for this stage of the analysis.

RESULTS

Extent of Disclosure

Our first purpose is to identify the extent to which our sample school districts used their web pages for financing referenda information disclosure in the March 2002 Illinois primary election. Overall, as indicated in Table 2, 43 of the 72 sample districts included at least some referenda-related information on their web sites. However, contrary to our hypotheses, there was little difference in the percentage of school districts making referenda disclosure across the referendum type and school district classification factors. The percentage of districts with referenda disclosure on their web pages was actually slightly higher for districts with bonds (61.5%) than for districts with tax referenda (57.6%). While urban and suburban school districts did have a higher percentage of disclosers than rural districts (63.5% versus 50.0%), neither this difference nor the one for the referendum type is statistically significant (based on a chi-square test for difference in the distribution). In contrast,

Table 2. The Use of Web Page Information Disclosure for a Sample of 72 Illinois School Districts with Financing Referenda on the March 19, 2002 Ballot.

Total sample	
School districts with referenda information disclosure	43 (59.7%)
School districts without referenda information disclosure	29 (40.3%)
By referendum type	
Building bond referenda	
School districts with referenda information disclosure	24 (61.5%)
School districts without referenda information disclosure	15 (38.5%)
Education fund tax referenda	
School districts with referenda information disclosure	19 (57.6%)
School districts without referenda information disclosure	14 (42.4%)
By school district classification	
Urban/suburban school districts	
School districts with referenda information disclosure	33 (63.5%)
School districts without referenda information disclosure	19 (36.5%)
Rural school districts	
School districts with referenda information disclosure	10 (50.0%)
School districts without referenda information disclosure	10 (50.0%)
By size	
Larger school districts (greater than 2,000 students)	
School districts with referenda information disclosure	23 (74.2%)
School districts without referenda information disclosure	8 (25.8%)
Smaller school districts (fewer that 2,000 students)	
School districts with referenda information disclosure	20 (48.8%)
School districts without referenda information disclosure	21 (51.2%)

and as expected, a significantly (at $p < 0.05$, two-tailed) higher percentage of larger school districts (74.2%) included web page referenda disclosure than did smaller districts (48.8%).

Table 3 presents information on the extensiveness of web page referenda disclosure by the sample school districts as based on the content analysis scores. Panel A of the table shows that the scores ranged from 0 to 9 (on a 10 point scale) with a mean score of 2.67. However, when only disclosing school districts are included, the mean disclosure score increases to 4.47. These results suggest that the school districts, on average, did not make particularly extensive financing referenda disclosures on their web sites.

Panel B of Table 3 provides data on the differences in the extent of disclosure across referendum type, school district classification, and district size factors. Similar to the results for the presence of disclosure, only size yields statistically

Table 3. Content Analysis Scores of the Extent of Web Page Referendum Disclosure by a Sample of 72 Illinois School Districts with Financing Referenda on the March 19, 2002 Ballot.

Panel A – Dispersion of content scores			
Minimum score	0	Maximum score	9
Mean score	2.67	Mean score (disclosers only)	4.47

	Mean Score (Std. Dev.)	Mann-Whitney *Z*-Statistic	Sig.[a]
Panel B – Differences in extent of disclosure			
By referendum type			
Building bond ($n = 39$)	2.56 (2.458)		
Education fund tax ($n = 33$)	2.79 (3.100)	0.117	0.907
By school district classification			
Urban/Suburban ($n = 52$)	2.75 (2.279)		
Rural ($n = 20$)	2.45 (2.874)	0.567	0.571
By school district size			
Larger ($n = 31$)	3.52 (2.920)		
Smaller ($n = 41$)	2.02 (2.465)	2.282	0.022

[a] Significance levels are two-tailed.

significant differences in the content disclosure scores. The mean content score for larger districts was 3.52 in comparison to a mean score of 2.02 for the smaller districts.

Relation to Election Outcome

Table 4 presents data on the percentage of successful referenda, broken down across web page disclosing and non-disclosing school districts. Overall, the sample of school districts including referenda information on their web pages exhibited a 74.4% pass rate in contrast to only a 44.8% pass rate for the sample of districts without web page referenda disclosure. This difference in pass rates is statistically significant at $p < 0.05$, two-tailed). However, the difference in pass rate percentages between districts with web page disclosure and without web page disclosure varies substantially when viewed across the referendum type, school district classification, and school district size factors. The pass rate percentage for districts including web page referenda disclosure is significantly higher (at at least $p < 0.05$, two tailed) for districts with education fund tax referenda, for urban and suburban districts, and for larger districts. In contrast, the pass rate percentage does not differ (statistically) across disclosers and non-disclosers

Table 4. Referenda Pass Rate Differences for a Sample of 72 Illinois School Districts with Financing Referenda on the March 19, 2002 Ballot Between Those Disclosing, and Those not Disclosing Referenda Information on Their Web Pages.

	Percentage of Passed Referenda[a]
Total sample	
School districts with web page disclosure	74.4
School districts with no web page disclosure	44.8*
By referendum type	
Building bond referenda	
School districts with web page disclosure	70.8
School districts with no web page disclosure	66.7
Education fund tax referenda	
School districts with web page disclosure	78.9
School districts with no web page disclosure	21.4**
By school district classification	
Urban/suburban	
School districts with web page disclosure	78.8
School districts with no web page disclosure	36.8**
Rural	
School districts with web page disclosure	60.0
School districts with no web page disclosure	60.0
By school district size	
Larger school districts (students >2000)	
School districts with web page disclosure	78.3
School districts with no web page disclosure	37.5*
Smaller school districts (students <2000)	
School districts with web page disclosure	70.0
School districts with no web page disclosure	47.6

[a] χ^2 tests were used to test for differences in the percentage of passed referenda across the various classifications.

* Significance in difference at the 0.05 level (two-tailed).

** Significance in difference at the 0.01 level (two-tailed).

for school districts with bond referenda, for rural school districts, or for smaller districts.

Results for the relation between disclosure content scores and election outcomes, reported in Table 5, mirror the results presented in Table 4. Results using a pass/no pass measure of election outcome (reported in Panel A) show a statistically significant (at $p < 0.01$ or better, two-tailed) correlation for: (1) the

Table 5. Correlations Between School District Web Page Disclosure Scores and Referenda Outcomes for a Sample of 72 Illinois School Districts with Financing Referenda on the March 19, 2002 Ballot.

	Correlation Coefficients	
	Parametric (Pearson *r*)	Non-Parametric (Spearman's rho)
Panel A – Correlation with election result (pass/no pass)		
Total sample ($n = 72$)	0.336**	0.339**
By referendum type		
Building bond ($n = 39$)	−0.074	−0.054
Education fund tax ($n = 33$)	0.714***	0.703***
By school district classification		
Urban/Suburban ($n = 52$)	0.491***	0.497***
Rural ($n = 20$)	−0.051	−0.038
By school district size		
Larger (n =31)	0.508**	0.513**
Smaller ($n = 41$)	0.171	0.192
Panel B – Correlation with percentage of "yes" votes		
Total sample ($n = 72$)	0.083	0.163
By referendum type		
Building bond ($n = 39$)	0.134	0.009
Education fund tax ($n = 33$)	0.382*	0.448**
By school district classification		
Urban/Suburban ($n = 52$)	0.283*	0.288*
Rural ($n = 20$)	–0.248	–0.142
By school district size		
Larger ($n = 31$)	0.303	0.362*
Smaller ($n = 41$)	0.011	0.119

*Statistical significance is designated by two-tailed (0.05 level).
**Statistical significance is designated by two-tailed (0.01 level).
***Statistical significance is designated by two-tailed (0.001 level).

overall sample; (2) districts with tax referenda; (3) urban/suburban districts; and (4) larger school districts. The relation between web page content scores and election outcome is not significant for districts with building bond referenda, for rural districts, or for the smaller school district sample.

In contrast to the results using the pass/no pass measure, the correlation between disclosure levels and the percentage of yes votes is not significant for the sample as a whole. However, with the exception that the parametric correlation measure is only significant at the $p = 0.097$ level for the larger school district breakdown,

results on sub-samples based on referenda type, school district classification, and school district size are consistent with the tests using the pass/no pass election result measure. That is, there is a significant relation between disclosure scores and election outcomes for districts with tax referenda, for urban/suburban districts, and for larger districts, but not for districts with bond referenda, for rural districts, or for smaller school districts. These correlations (and their levels of significance), however, are not as strong as their counterparts using the pass/no pass measure.

DISCUSSION AND LIMITATIONS

Bowler and Donovan (1998), building on Downs' (1957) model of election choice, assert that a major focus in direct democracy campaigns should be to provide better information to the electorate. Accordingly, the first purpose of our analysis is to identify the extent to which school districts are using the Internet, a potentially powerful communications tool, for providing information on financing referenda. From a purely descriptive perspective, our review indicates that school districts, at least in Illinois, are beginning to use their web sites for referenda information disclosure. More than half of the sample districts with referenda on the March 19, 2002 election ballot included at least some disclosure. However, as evidenced by the relatively low mean content scores, it appears that there is considerable opportunity for even greater web page information dissemination. Given the lack of mention of the Internet in the earlier referenda guides (Boschee & Holt, 1999; Graves, 1999), it is perhaps not surprising that web page disclosure is not yet extensive. Identifying whether Internet use for information dissemination by school districts increases in future years would appear to be a potentially interesting extension of this research. A further valuable extension would be to examine the extent to which school districts bias their information disclosures to voters in order to increase the likelihood of successful passage.

A second focus of our analysis is to identify whether school districts' use of their web pages for referenda disclosure varies across systematic factors. In general, the answer appears to be that it does not. Districts with bond referenda on the ballot were as likely to include web page information disclosure as districts with tax referenda, and rural school districts were as likely as urban and suburban school districts to make referenda disclosure on their web sites. Similarly, there were no statistical differences in the extensiveness of the disclosure across either of these factor classifications. The only significant difference in the use of web pages for referenda disclosure across categories was relative to school district

size. As hypothesized, larger districts were both more likely to include disclosure and to have greater levels of disclosure than smaller school districts.

Finally, beyond just examining the extent of web page financing referenda information disclosure by school districts, we also examine whether the disclosure is associated with the election outcome. Both Popkin (1993) and Lupia (1994) suggest that because voters rely on shortcuts and other information cues, the value of the Internet as an information disclosure medium might be expected to vary. Consistent with this argument, we find significant differences in the relation between Internet disclosure and referenda success across school district classifications. Greater web page disclosure was correlated with higher referenda success (both in terms of pass/no pass classification and the percentage of yes votes) for districts with tax as opposed to bond referenda, for urban or suburban districts as opposed to rural districts, and for larger as opposed to smaller school districts. This suggests that the value to using the web for referenda information dissemination, as expected, may vary across school districts.

There are at least two major limitations to our analysis that force us to suggest using caution in interpreting our results with respect to disclosure and its relation to election outcome. First, it is very possible that school districts that chose to include greater levels of referenda disclosure on their web pages also exhibited higher levels of information disclosure through other, more traditional means (e.g. flyers, newspaper ads, informational meetings, etc.). If true, it may be the overall level of information disclosure (and its correlation to web page disclosure) that is driving the relation with election outcome. Future analysis should attempt to control for this possibility by surveying school district administrators to generate some measure of overall information disclosure. However, if it is true that our web disclosure measures correlate with an overall information disclosure strategy, it would appear that more research into the costs vs. the benefits of greater disclosure across differing types of school districts with differing types of financing referenda on the ballot is warranted.

The second limitation to our analysis is our failure to control for other factors that may affect the referenda election results. Lentz (1999), for example, notes that factors including community educational culture, capacity to pay, the existence of formal opposition groups, and voter turnout, among others, have all been shown to influence referenda outcomes in at least some situations. However, as noted by both Hamilton and Cohen (1974) and Freeman (1990), none of these factors are consistently related to referenda success or failure. To more definitively ascribe a relation between web page information disclosure and referenda outcome, these, and perhaps other factors should be addressed. However, and importantly, we believe our exploratory findings provide support for pursuing such future investigations.

NOTES

1. Data from the Illinois State Board of Education indicates a similar pattern of defeat for school district tax referenda over the past five years.

2. Bowler and Donovan (1998, p. 69) note that these expectations with respect to *no* voting follow not only from conventional wisdom, but also from a large body of research in political science and psychology on what they term "negativity bias."

3. In defense of Boschee and Holt (1999), they appear to rely on a 1994 American Association of School Administrators study that examined where voters claimed to get their information. Web use at that time was significantly less prominent than it is today.

4. As noted below, our findings suggest that Internet disclosure may not in fact be "essential" for all districts.

5. A review of both practice and academic journals in the Public Administration, Governmental Accounting, and School Administration literature for the past three years revealed no articles reporting on Internet usage by school districts.

6. Chan and Rubin (1987, p. 4) argue that a major goal of governmental accounting research "is to describe, explain, or predict how information . . . is used by decision makers in making economic, political and social choices in the political system and government." Sanders (1994, p. 138) further argues that accounting researchers might add to the body of knowledge by identifying how information impacts political decisions, much as they have previously examined the role of information in valuing investments and in contracting situations. Examining whether Internet access to detailed referenda data correlates with voter decisions would appear to fit nicely with the calls of both Chan and Rubin (1987) and Sanders (1994).

7. Evidence from recent elections in Minnesota and Wisconsin suggest similar patterns of referenda success and failure in those states.

8. According to the ISBE (at http://www.isbe.net), classification is based on a combination of three criteria. These are: (1) the proportions of each school district's population classified as urban residents and rural residents, respectively, by the U.S. Bureau of the Census; (2) the location of the school district, whether inside or outside of a metropolitan area; and (3) the area served primarily by the school district, whether inside or outside the central city/cities of a metropolitan area.

9. There is also at least some evidence that metropolitan-based school districts have a harder time passing financing referenda. For example, based on information from Education Minnesota (http://www.educationminnesota.org) and the Taxpayers League of Minnesota (http://www.taxpayersleague.com), while the overall pass rate for school district financing referenda in the November 2001 election was 68.5%, the state's metropolitan-based districts experienced only a 31% pass rate.

10. The results reported here for differences across school district size were based on a definition of "larger" school districts as those with more than 2,000 students and "smaller" school districts as those with fewer than 2,000 students. Additional tests using both alternative "cut-off" points, and deleting various numbers of the middle size school districts provided results that were qualitatively similar to those reported here. Further, correlation analysis (both parametric and non-parametric) indicated a statistically significant (at $p < 0.05$, two-tailed) relation between the number of students and school district web page disclosure scores.

11. As noted by an anonymous reviewer, another reason school district size may be a factor is that web costs are relatively fixed regardless of school district size. Thus, given presumably larger budgets, it may be easier for larger school districts to underwrite the cost of a web presence.

12. According to the State of Illinois' 2002 Election and Campaign Finance Calendar, all referenda to be placed on the March 19, 2002 ballot had to be certified by the appropriate county clerk's office no later than January 17, 2002.

13. This district, Round Lake Area District #116, had an education fund tax levy referendum on the ballot. It was defeated. Because we had not identified this district prior to the election we were unable to determine whether it included referenda disclosure on its web page and it was not included in any of our analyses. We cannot assess the extent to which this omission may impact our reported findings.

14. Our concern is to eliminate other potential factors that might explain differences in school district web page disclosure of referenda information and its relation to election results.

15. We repeated all tests coding these 11 districts as "zero" disclosers. Results, in all cases, were qualitatively similar to those reported.

16. More specifically, one of the project co-authors conducted telephone interviews with school district administrators from two Illinois districts that had successfully passed financing referenda (one tax referenda and one bond referenda) in the recent past.

17. Numerous studies (see, e.g. Wiseman, 1982) use weighted coding schemes in their content analysis of disclosures. To assure that the results we report are not a function of this scheme we repeated all tests using a six point scale where no weighting was used for any disclosure areas. Results for all tests were qualitatively similar to those reported in the paper.

18. While there were numerous districts that added disclosure between the two reviews, we found no case where disclosure was reduced.

REFERENCES

Ashbaugh, H., Johnstone, M., & Warfield, T. D. (1999). Corporate reporting on the internet. *Accounting Horizons* (September), 241–258.

Balutis, A. P. (2001). E-Government, Part II: Evolving strategies for action. *The Public Manager* (Summer), 41–45.

Boer, H. (2001). *Illinois school district referenda*. Macomb, IL: Illinois Institute for Rural Affairs, Western Illinois University.

Boschee, F., & Holt, C. R. (1999). *School bond success*. Lancaster, PA: Technomic Publishing.

Bowler, S., & Donovan, T. (1998). *Demanding choices*. Ann Arbor: University of Michigan Press.

Chan, J. L., & Rubin, M. A. (1987). The role of information in a democracy and in government operations: The public choice methodology. *Research in Governmental and Nonprofit Accounting*, 3–27.

Cohen, C. (1971). *Democracy*. Athens, GA: University of Georgia Press.

Dewees, S. (1999). Improving rural school facilities for teaching and learning. *ERIC Digests* (December). Available at www.ericfacility.net/databases/ERIC_Digests/index/.

Downs, A. (1957). *An economic theory of democracy*. New York: Harper & Row.

Freeman, R. E. (1990). *Perceptions of Illinois school superintendents concerning what variables predict success in educational fund referendum campaigns*. Unpublished dissertation, St. Louis University.

Graves, J. J. Jr. (1999). *Winning the referendum battle in education*. Glen Ellyn, IL: Jomil Publishing.

Hamilton, H. D., & Cohen, S. H. (1974). *Policy making by Plebiscite: School referenda*. Lexington, MA: D. C. Heath and Company.

Hammel, D. J. (2001). *Internet use among local governments in Illinois' nonmetropolitan communities*. Macomb, IL: Illinois Institute for Rural Affairs, Western Illinois University.

Hobbs, D. (1995). Capacity building: Reexamining the role of the rural school. In: L. J. Beaulieu & D. Mulkey (Eds), *Investing in People: The Human Capital Needs of Rural America* (Ch. 12). Boulder, CO: Westview Press.

Jones, K., Alabaster, T., & Hetherington, K. (2000). Internet-based environmental reporting. *GMI* (Summer), 69–90.

Lally, L. (1996). Privacy vs. accountability: The impact of situationally conditioned belief. *Journal of Business Ethics* (November), 1221–1226.

Lentz, C. (1999). Predicting school referenda outcomes: Answers from Illinois. *Journal of Education Finance* (Spring), 459–482.

Lupia, A. (1994). Shortcuts vs. encyclopedias: Information and voting behavior in California insurance reform elections. *American Political Science Review* (March), 63–76.

Magleby, D. B. (1984). *Direct legislation: Voting on ballot propositions in the U.S.* Baltimore: Johns Hopkins University Press.

Moulder, E. (2001). E-Government . . . If you build it, will they come? *Public Management* (September), 10–14.

Norris, D., Fletcher, P., & Holden, S. (2001). Is your local government plugged in? *Public Management* (June), 4–10.

Patten, D. M. (2002). Give or take on the internet: An examination of the disclosure practices of insurance firm web innovators. *Journal of Business Ethics* (March), 247–259.

Patten, D. M., & Crampton, W. (2004). Legitimacy and the internet: An examination of corporate web page environmental disclosures. *Advances in Environmental Accounting and Mangement*, 31–57.

Popkin, S. L. (1991). *The reasoning voter*. Chicago: University of Chicago Press.

Popkin, S. L. (1993). Information shortcuts and the reasoning voter. In: B. Grofman (Ed.), *Information, Participation, and Choice* (Ch. 1). Ann Arbor: University of Michigan Press.

Rado, D. (2002). Deficits rack school districts. *Chicago Tribune* (2 January), 1, 8.

Sanders, G. D. (1994). Tiebout migrants and the median voter as users of accounting information. *Research in Governmental and Nonprofit Accounting*, 137–155.

Wildstrom, S. H. (1997). Surfing for annual reports. *Business Week* (14 April), 22.

Wiseman, J. (1982). An evaluation of environmental disclosures made in corporate annual reports. *Accounting, Organizations and Society*, 53–63.

APPENDIX

Content Analysis Disclosure Items

	Number of Districts Disclosing
Tax referenda disclosure items	
Narrative justification for the Levy (3 point scale)	
1 – Limited narrative justification	6
2 – Moderate narrative justification	4
3 – Extensive narrative justification	8
History of past requests or increases (1 point scale)	14
Comparative tax rate data (2 point scale)	
1 – Either multi-year comparisons or comparisons to other districts	4
2 – Both multi-year and other districts comparisons	1
Impact of Levy on taxpayer (payment amounts) (1 point scale)	11
Use of budget or actual financial data (2 point scale)	
1 – Summary financial data only	5
2 – Detailed financial data	8
Sample ballot provided (1 point scale)	2
Total potential disclosure score = 10	
Bond referenda disclosure items	
Narrative justification for bond issuance (3 point scale)	
1 – Limited narrative justification	10
2 – Moderate narrative justification	7
3 – Extensive narrative justification	7
History of past funding requests or increases (1 point scale)	4
Project cost estimates (1 point scale)	13
Architect's drawings/blueprints (1 point scale)	12
Impact of bonds on taxpayer (payment amounts) (1 point scale)	19
Use of budget or actual financial data (2 point scale)	
1 – Summary data only	4
2 – Detailed data	1
Sample ballot provided (1 point scale)	7
Total potential disclosure score = 10	