Volume 12, Number 3 **ISSN 1096-3685**

ACADEMY OF ACCOUNTING AND FINANCIAL STUDIES JOURNAL

Michael Grayson, Jackson State University
Accounting Editor

Denise Woodbury, Southern Utah University
Finance Editor

Academy Information
is published on the Allied Academies web page
www.alliedacademies.org

The *Academy of Accounting and Financial Studies Journal* is owned and published by the DreamCatchers Group, LLC, and printed by Whitney Press, Inc. Editorial content is under the control of the Allied Academies, Inc., a non-profit association of scholars, whose purpose is to support and encourage research and the sharing and exchange of ideas and insights throughout the world.

Whitney Press, Inc.

Printed by Whitney Press, Inc.
PO Box 1064, Cullowhee, NC 28723
www.whitneypress.com

Authors execute a publication permission agreement and assume all liabilities. Neither the DreamCatchers Group or Allied Academies is responsible for the content of the individual manuscripts. Any omissions or errors are the sole responsibility of the authors. The Editorial Board is responsible for the selection of manuscripts for publication from among those submitted for consideration. The Publishers accept final manuscripts in digital form and make adjustments solely for the purposes of pagination and organization.

The *Academy of Accounting and Financial Studies Journal* is owned and published by the DreamCatchers Group, LLC, PO Box 2689, 145 Travis Road, Cullowhee, NC 28723. Those interested in subscribing to the *Journal*, advertising in the *Journal*, submitting manuscripts to the Journal, or otherwise communicating with the *Journal,* should contact the Executive Director at info@alliedacademies.org.

Copyright 2008 by the DreamCatchers Group, LLC, Cullowhee, NC, USA

Academy of Accounting and Financial Studies Journal
Accounting Editorial Review Board Members

Agu Ananaba
Atlanta Metropolitan College
Atlanta, Georgia

Manoj Anand
Indian Institute of Management
Pigdamber, Rau, India

Ali Azad
United Arab Emirates University
United Arab Emirates

D'Arcy Becker
University of Wisconsin - Eau Claire
Eau Claire, Wisconsin

Jan Bell
California State University, Northridge
Northridge, California

Linda Bressler
University of Houston-Downtown
Houston, Texas

Jim Bush
Middle Tennessee State University
Murfreesboro, Tennessee

Douglass Cagwin
Lander University
Greenwood, South Carolina

Richard A.L. Caldarola
Troy State University
Atlanta, Georgia

Eugene Calvasina
Southern University and A & M College
Baton Rouge, Louisiana

Darla F. Chisholm
Sam Houston State University
Huntsville, Texas

Askar Choudhury
Illinois State University
Normal, Illinois

Natalie Tatiana Churyk
Northern Illinois University
DeKalb, Illinois

Prakash Dheeriya
California State University-Dominguez Hills
Dominguez Hills, California

Rafik Z. Elias
California State University, Los Angeles
Los Angeles, California

Richard Fern
Eastern Kentucky University
Richmond, Kentucky

Peter Frischmann
Idaho State University
Pocatello, Idaho

Farrell Gean
Pepperdine University
Malibu, California

Luis Gillman
Aerospeed
Johannesburg, South Africa

Richard B. Griffin
The University of Tennessee at Martin
Martin, Tennessee

Marek Gruszczynski
Warsaw School of Economics
Warsaw, Poland

Morsheda Hassan
Grambling State University
Grambling, Louisiana

Richard T. Henage
Utah Valley State College
Orem, Utah

Rodger Holland
Georgia College & State University
Milledgeville, Georgia

Kathy Hsu
University of Louisiana at Lafayette
Lafayette, Louisiana

Shaio Yan Huang
Feng Chia University
China

Robyn Hulsart
Ohio Dominican University
Columbus, Ohio

Evelyn C. Hume
Longwood University
Farmville, Virginia

Terrance Jalbert
University of Hawaii at Hilo
Hilo, Hawaii

Marianne James
California State University, Los Angeles
Los Angeles, California

Academy of Accounting and Financial Studies Journal
Accounting Editorial Review Board Members

Jongdae Jin
University of Maryland-Eastern Shore
Princess Anne, Maryland

Ravi Kamath
Cleveland State University
Cleveland, Ohio

Marla Kraut
University of Idaho
Moscow, Idaho

Jayesh Kumar
Xavier Institute of Management
Bhubaneswar, India

Brian Lee
Indiana University Kokomo
Kokomo, Indiana

Harold Little
Western Kentucky University
Bowling Green, Kentucky

C. Angela Letourneau
Winthrop University
Rock Hill, South Carolina

Treba Marsh
Stephen F. Austin State University
Nacogdoches, Texas

Richard Mason
University of Nevada, Reno
Reno, Nevada

Richard Mautz
North Carolina A&T State University
Greensboro, North Carolina

Rasheed Mblakpo
Lagos State University
Lagos, Nigeria

Nancy Meade
Seattle Pacific University
Seattle, Washington

Thomas Pressly
Indiana University of Pennsylvania
Indiana, Pennsylvania

Hema Rao
SUNY-Oswego
Oswego, New York

Ida Robinson-Backmon
University of Baltimore
Baltimore, Maryland

P.N. Saksena
Indiana University South Bend
South Bend, Indiana

Martha Sale
Sam Houston State University
Huntsville, Texas

Milind Sathye
University of Canberra
Canberra, Australia

Junaid M.Shaikh
Curtin University of Technology
Malaysia

Ron Stunda
Birmingham-Southern College
Birmingham, Alabama

Darshan Wadhwa
University of Houston-Downtown
Houston, Texas

Dan Ward
University of Louisiana at Lafayette
Lafayette, Louisiana

Suzanne Pinac Ward
University of Louisiana at Lafayette
Lafayette, Louisiana

Michael Watters
Henderson State University
Arkadelphia, Arkansas

Clark M. Wheatley
Florida International University
Miami, Florida

Barry H. Williams
King's College
Wilkes-Barre, Pennsylvania

Carl N. Wright
Virginia State University
Petersburg, Virginia

Academy of Accounting and Financial Studies Journal
Finance Editorial Review Board Members

Confidence W. Amadi
Florida A&M University
Tallahassee, Florida

Roger J. Best
Central Missouri State University
Warrensburg, Missouri

Donald J. Brown
Sam Houston State University
Huntsville, Texas

Richard A.L. Caldarola
Troy State University
Atlanta, Georgia

Darla F. Chisholm
Sam Houston State University
Huntsville, Texas

Askar Choudhury
Illinois State University
Normal, Illinois

Prakash Dheeriya
California State University-Dominguez Hills
Dominguez Hills, California

Martine Duchatelet
Barry University
Miami, Florida

Stephen T. Evans
Southern Utah University
Cedar City, Utah

William Forbes
University of Glasgow
Glasgow, Scotland

Robert Graber
University of Arkansas - Monticello
Monticello, Arkansas

John D. Groesbeck
Southern Utah University
Cedar City, Utah

Marek Gruszczynski
Warsaw School of Economics
Warsaw, Poland

Mahmoud Haj
Grambling State University
Grambling, Louisiana

Mohammed Ashraful Haque
Texas A&M University-Texarkana
Texarkana, Texas

Terrance Jalbert
University of Hawaii at Hilo
Hilo, Hawaii

Ravi Kamath
Cleveland State University
Cleveland, Ohio

Jayesh Kumar
Indira Gandhi Institute of Development Research
India

William Laing
Anderson College
Anderson, South Carolina

Helen Lange
Macquarie University
North Ryde, Australia

Malek Lashgari
University of Hartford
West Hartford, Connetticut

Patricia Lobingier
George Mason University
Fairfax, Virginia

Ming-Ming Lai
Multimedia University
Malaysia

Steve Moss
Georgia Southern University
Statesboro, Georgia

Christopher Ngassam
Virginia State University
Petersburg, Virginia

Bin Peng
Nanjing University of Science and Technology
Nanjing, P.R.China

Hema Rao
SUNY-Oswego
Oswego, New York

Milind Sathye
University of Canberra
Canberra, Australia

Daniel L. Tompkins
Niagara University
Niagara, New York

Randall Valentine
University of Montevallo
Pelham, Alabama

Marsha Weber
Minnesota State University Moorhead
Moorhead, Minnesota

ACADEMY OF ACCOUNTING AND FINANCIAL STUDIES JOURNAL

CONTENTS

LETTER FROM THE EDITORS

Welcome to the *Academy of Accounting and Financial Studies Journal.* The editorial content of this journal is under the control of the Allied Academies, Inc., a non profit association of scholars whose purpose is to encourage and support the advancement and exchange of knowledge, understanding and teaching throughout the world. The mission of the *AAFSJ* is to publish theoretical and empirical research which can advance the literatures of accountancy and finance.

Dr. Michael Grayson, Jackson State University, is the Accountancy Editor and Dr. Denise Woodbury, Southern Utah University, is the Finance Editor. Their joint mission is to make the *AAFSJ* better known and more widely read.

As has been the case with the previous issues of the *AAFSJ*, the articles contained in this volume have been double blind refereed. The acceptance rate for manuscripts in this issue, 25%, conforms to our editorial policies.

The Editors work to foster a supportive, mentoring effort on the part of the referees which will result in encouraging and supporting writers. They will continue to welcome different viewpoints because in differences we find learning; in differences we develop understanding; in differences we gain knowledge and in differences we develop the discipline into a more comprehensive, less esoteric, and dynamic metier.

Information about the Allied Academies, the *AAFSJ*, and our other journals is published on our web site. In addition, we keep the web site updated with the latest activities of the organization. Please visit our site and know that we welcome hearing from you at any time.

Michael Grayson, Jackson State University

Denise Woodbury, Southern Utah University

www.alliedacademies.org

FACTORS ASSOCIATED WITH THE LEVEL OF SUPERFUND LIABILITY DISCLOSURE IN 10K REPORTS: 1991-1997

Carol A. Cox, Middle Tennessee State University

ABSTRACT

This study examines factors associated with the level of Superfund disclosure in 10K reports. Sample firms consist of Fortune 500 companies identified by the Environmental Protection Agency as Potentially Responsible Parties. The study utilizes a comprehensive environmental disclosure index based on Generally Accepted Accounting Principles, to measure the extent to which sample firms disclose environmental liability information. Empirical tests demonstrate that the extent of environmental disclosure is associated with size, profitability, industry classification and regulatory influence.

The study uses data from COMPUSTAT, EDGAR, and the Superfund Public Information System for years 1991-1997. The environmental disclosure index is compiled based on relevant authoritative guidance contained in Regulation S-K, SAB 92, and SFAS 5. Policy implications indicate that the Securities and Exchange Commission must improve monitoring and enforcement efforts designed to promote adequate recognition and disclosure related to environmental liabilities.

INTRODUCTION

Over the past fifteen years, there has been increased attention on the environment and how companies measure and report environmental exposure (Cox, 2001). Compliance with environmental regulations has a significant impact on corporate earnings, particularly for certain industries. Fitzgerald (1996) estimates that in excess of $7 billion a year is spent for U.S. Superfund site remediation. The term "Superfund" refers to the federal trust fund established to pay for environmental cleanup and enforcement. Furthermore, between $500 and $750 billion will be required to remediate all sites identified by the Environmental Protection Agency (EPA) under the Comprehensive Environmental Response, Compensation and Liability Act (the Superfund Act) (Lavelle, 1992; Committee on Commerce 1995). Probst et al. (1995) estimate that $6 billion is spent each year pursuant to the Superfund Act and $135 billion is required annually to comply with all federal environmental regulations. In light of the magnitude of estimated costs, the reporting of environmental costs and obligations has become a prominent issue with accounting regulatory and professional bodies, such as the Securities Exchange Commission (SEC), Financial Accounting Standards Board (FASB), American Accounting Association (AAA) and American Institute of

Certified Public Accountants (AICPA) (Sack et al, 1995). Of specific concern are the recognition of environmental liabilities and the adequacy of related environmental disclosures (ED).

Recent accounting scandals have resulted in increased public scrutiny of corporate governance and disclosures, and have further heightened the attention on ED (Bibler, 2003). The new regulatory requirements promulgated by the Sarbanes-Oxley Act of 2002 have implications for companies with environmental obligations. United States Senators have asked the Government Accounting Office to examine whether SEC requirements for ED are adequate (Sissell, 2002). Moreover, philanthropic foundations and investment mangers are appealing to the SEC to improve enforcement activities related to ED (Bank, 2002).

Early studies suggest that ED are self-serving and inaccurate, although much of the evidence is anecdotal [Beams and Fertig, 1971; Estes, 1976; Churchill, 1978; Nader, 1978]. A later stream of empirical research, resulting from growing public scrutiny of ED, examines both the content and quality of environmental disclosure (Gamble et al., 1995; Freedman and Wasley, 1990; Rockness, 1985; Wiseman, 1982, Ingram and Frazier, 1980). This research suggests that ED quality is generally low because the disclosures do not adequately reflect the firm's actual environmental liability exposure. In addition, firms generally do not record material environmental liabilities. The public scrutiny of ED as well as the findings of early research led to additional reporting requirements with respect to environmental liabilities, as well as increased oversight of environmental matters by the SEC. Two later studies suggest that firms have increased the extent of ED in response to Staff Accounting Bulletin No. 92 (SAB 92, issued in 1993); however, they have not increased disclosure of accrued amounts for environmental remediation (Stanny, 1998; Barth et al., 1997). Thus, like previous studies, Stanny and Barth et al. conclude that accounting guidance is not successful in promoting the recognition of environmental liabilities.

The current study maintains that while much discretion exists under Generally Accepted Accounting Principles (GAAP), firms are subject to specific, mandatory recognition and disclosure requirements with respect to environmental liabilities. Thus, the current study compiles a comprehensive index of GAAP disclosures related to environmental liabilities, based on relevant authoritative guidance provided by the FASB and SEC. The index is used to measure the extent of ED in the 10K filings of 182 Fortune 500 companies with known potential environmental liabilities, for years 1991-1997. The research question is: Are firm characteristics associated with the amount of ED included in 10K reports? The study examines whether size, industry classification, profitability, and regulatory influence affects the extent of ED provided in 10K reports. Prior research finds the level of social disclosure to be associated with industry and size, but not profitability (Trotman and Bradley, 1981; Cowen et al., 1987; Patten, 1991). Stanny (1998) and Barth et al. (1997) suggest that increased regulation is associated with the level of ED. The current study finds that size, industry classification, profitability, and regulatory influence are significantly associated with the level of ED in 10K reports.

This study contributes to the understanding of ED practices of publicly traded U.S. firms because it measures the extent of ED included in 10K filings using a comprehensive index of

relevant GAAP, and it identifies factors that influence the level of ED. Understanding the disclosure practices of publicly traded firms allows investors, creditors, regulators and standard setters to determine the adequacy of ED, and to assess the need for additional reporting guidance. The study contributes to recent discussion between regulators and standard-setters regarding the adequacy of U.S. accounting standards, and how FASB guidance inhibits transparency, which is central to the SEC's goal of full and fair disclosure (Herdman, 2002).

Understanding factors that influence the level of ED is important to researchers that seek to develop a theory of social responsibility disclosure. Prior studies show ED varies greatly with respect to the quantity and quality of information provided (Price Waterhouse, 1992, 1994; Gamble et al., 1995; Kreuze, 1996). Spicer (1978) suggests that a convenient starting point in theory building is the observation and description of the real world and noting correlations between variables of interest. Therefore, several studies investigate characteristics of companies that may be associated with their social responsibility disclosures (Trotman and Bradley, 1981; Cowen et al., 1987; Patten, 1991). The current study provides evidence regarding the association between four independent variables (size, industry, profitability, and regulatory influence) and the level of ED.

The remainder of this paper is presented as follows. Section 2 provides institutional background, summarizing environmental regulation and reporting requirements. Section 3 reviews the relevant ED literature. Section 4 presents the research methodology; including sample selection procedures, model variable descriptions, and hypotheses. Section 5 presents the empirical results and analysis. Section 6 provides discussion and conclusions.

INSTITUTIONAL BACKGROUND

Statement of Position 96-1, Environmental Remediation Liabilities, summarizes the regulatory process with regard to the Superfund Act (AICPA, 1996). The Superfund Act regulates the cleanup of inactive waste disposal sites and spills. The Superfund Act adopted a "polluter pays" philosophy by establishing the right to bill firms associated with sites for their portion of the remediation costs and levying a tax on certain industries to fund orphan sites. Several features of the Superfund Act present challenges for estimating a firm's liability under its provisions. First, the liability is for response and remediation costs, as well as for damages and health assessment and study costs. Depending on the nature of the contamination and the cleanup technology chosen, remediation costs can include initial capital expenditures in the millions of dollars and ongoing operating, maintenance, and monitoring costs for 30 or more years. Moreover, cleanup standards are unspecified. Second, the Superfund Act imposes liability on a broad group of potentially responsible parties (PRPs) that includes the site's current owner, and anyone who: owned or operated the facility when hazardous substances were disposed, generated hazardous substances disposed of at the facility, transported hazardous substances to the disposal facility, and/or arranged for such transportation. Third, the Superfund Act liability is strict, retroactive, and joint and several. Strict liability means that the party is liable without regard to fault. The EPA need not prove negligence.

The law is retroactive in that the liability is imposed for actions that may not have violated the law at the time. Under joint and several liability, which is permitted but not required by Section 9607 of the Superfund Act, any PRP can be held responsible for the full cost of cleanup if the harm is indivisible. This means that the "deep pocketed" party is often legally liable for all remediation costs. Other PRPs, if available, could be sued subsequently for their share of cleanup costs. Environmental obligations arising under the Superfund Act are the prime focus of this paper because they include some of the largest cleanup obligations of publicly traded corporations.

GAAP provides specific recognition and disclosure requirements with respect to environmental liabilities. The most relevant GAAP related to environmental liabilities is Statement of Financial Accounting Standards Number 5 (SFAS 5), Accounting for Contingencies (FASB, 1975), which establishes both recognition rules and disclosure rules for contingent liabilities. SFAS 5 states that a loss contingency must be accrued if it is both probable and reasonably estimable. In the context of environmental obligations, the first condition is often met when the EPA notifies the organization of its status as a PRP. The second condition, however, is more difficult to determine. If no accrual is made because one or both conditions are not met, the standard provides disclosure rules for reasonably possible losses. The disclosure should indicate the nature of the contingency and should give an estimate of the possible loss, range of possible losses or state that such an estimate cannot be made.

In addition to the requirements of SFAS 5, firms must comply with guidance issued by the SEC. Most relevant to the current study is Regulation S-K (revised in 1986): Items 101, 103, and 303 (SEC, 2000), and Staff Accounting Bulletin 92 (SAB 92) (SEC, 1993). Regulation S-K provides standard instructions for filing forms under the Securities Act of 1933, Securities Act of 1934, and Energy Policy and Conservation Act of 1935. It describes several items to be included in 10K filings. Items 101, 103, and 303 provide general guidance for disclosure of environmental information in the 10K. Item 101 requires a general description of the business and specific disclosure of the effects that compliance with environmental laws may have on capital expenditures, earnings, and competitive position, when material. In addition, the estimated amount disclosed for capital expenditures should apply to the current and succeeding fiscal years and any future periods in which those expenditures may be material. Item 103 requires disclosure of pending or contemplated administrative or judicial proceedings, whereas, Item 303 requires disclosure of material events and uncertainties known to management that would cause reported financial information to be unrepresentative of future operating results or financial conditions. However, SAB 92 was issued specifically to improve the disclosure of environmental liability information. In particular, it makes it inappropriate to present environmental liabilities net of claims for insurance recovery. SAB 92 also requires additional disclosures related to discounting of liabilities and insurance recoveries.

SFAS 5, Regulation S-K and SAB 92 provide ED requirements for firms with known environmental obligations. Based on this authoritative guidance, the current study utilizes a

comprehensive ED index to measure the extent of ED in the 10K reports of sample firms. The resulting measure provides the basis for the examination of factors associated with the level of ED.

RELATED LITERATURE

Studies that investigate the amount and content of ED in annual reports and 10Ks suggest that ED has been largely voluntary (Berthelot et al., 2003). These studies reveal a wide variety of disclosure practices by firms with respect to disclosure quantity and quality. Researchers seek to determine what environmental information companies disclose, and whether or not such disclosures are adequate to meet the needs of various stakeholders. Through descriptive and survey methodologies, prior research suggests that ED quality is generally low, and that firms generally do not record environmental liabilities (Price Waterhouse, 1992, 1994; Gamble et al., 1995; Kreuze, 1996; Walden et al., 1997). Researchers in this area suggest these results are due to vague or inadequate reporting standards with respect to environmental information. The studies, however, do not attempt to determine whether environmental disclosures are consistent with GAAP requirements, thereby utilizing a comprehensive ED index. In addition, the studies are conducted in a vastly different regulatory environment—prior to SAB 92 and increased SEC oversight. Moreover, the studies do not utilize EPA liability data to determine the potential liability exposure of the sample firms. Thus, the current study contributes to this area of research by examining the disclosure practices of firms with known environmental liability exposure, as determined by the EPA. Also, the current study focuses on ED outlined by GAAP, in order to determine the level of GAAP disclosure provided over time.

The current study relates most closely to two prior studies that examine the level of ED in response to SAB 92. Using a sample of 199 firms, Stanny (1998) finds significant increases in the reporting of eight environmental disclosures during the period from 1991-1993. However, because the average per firm increase in accrued liability is an insignificant percentage of market value of equity, the author concludes that SAB 92 has had limited success in improving recognition of environmental liabilities. Using a sample of 257 firms in the auto, chemical, appliances, and utilities industries, Barth et al. (1997) find that regulatory influence is associated with firms' environmental disclosure decisions, and firms with larger estimated liabilities. Additionally, the researchers report increased disclosure over the sample period from 1989-1993. Thus, they conclude that accounting regulations have an effect on environmental disclosure. However, sample firms accrue an environmental liability in only 34% of the sample firm-years. As a result, Barth et al. conclude that firms have considerable discretion with respect to environmental liability disclosures.

The current study differs from previous studies in several ways. Most importantly, the level of ED is measured based on a comprehensive disclosure index of GAAP requirements. In addition, the sample is limited to firms with known potential environmental obligations. Proxies for environmental liability are based directly on cost estimates provided by the EPA in its Record of Decision (ROD). The ROD contains the first publicly available cleanup cost estimates, and it is

completed after an extensive EPA investigation. The current study also examines a longer and more recent sample period (1991-1997) than previous research on ED. Moreover, in addition to examining the impact of regulation on ED, the current study examines whether firm characteristics such as size, industry and profitability affect the extent of ED.

RESEARCH METHOD

The sample is taken from the 1997 Fortune 500, which represents companies with the highest revenues, or "the deepest pockets". The sampling procedures are designed to obtain a sample of public U.S. firms with known estimated environmental liabilities (which proxies for regulatory influence), and the ability to pay (as remediation is most often paid by the deep-pocketed firms). Thus, the following conditions must be met for inclusion in the sample: 1) the company must be named as a PRP on at least one Superfund Site throughout the entire sample period (1991-1997), 2) the company must be non-financial and publicly traded, and 3) form 10K data must be available from 1991-1997. Non-financial refers to firms not classified in Standard Industrial Classification (SIC) Division H (finance, insurance). Due to the nature of the business, financial institutions have no or limited environmental exposure and are therefore excluded from the current sample. Because the Fortune 500 list includes companies that must report part or all of their figures to a government agency, private companies that produce a 10K are included. The current study is interested in publicly traded companies, and therefore excludes private companies

Table 1 summarizes the sample selection procedures. The initial sample consists of 245 firms named as a PRP on at least one site in the Superfund Public Information System (SPIS) database. The Standard Industrial Classification (SIC) is then obtained for each sample firm from the Lexus Nexus database. Firms in SIC Division H (finance, insurance) are excluded from the sample. Twenty-seven firms are named to a site not included on the final National Priorities Listing (NPL) and were not issued a Record of Decision (ROD). The Edgar database is used to obtain 10K data for all sample years. Eleven firms did not file 10Ks during the entire sample period from 1991-1997, and are excluded from the sample. Thirteen firms are named as PRPs during the sample period, and therefore do not have liabilities for the entire sample period. The final sample consists of 182 firms in 33 industries. Financial data for sample firms is obtained from 10K filings and from

COMPUSTAT

For the dependent variable, the current study utilizes a comprehensive listing of environmental liability disclosures (ED Index), compiled based on Regulation S-K (items 101, 103 and 303), SAB 92, and SFAS 5. Table 2 summarizes the twenty-nine disclosure items. Firm 10K reports are examined for the presence or absence of specific statements as outlined in the ED Index (for fiscal years 1991-1997). Two reviewers (the author and a research assistant) evaluate each 10K

report independently. The reviewers met weekly to discuss independent evaluations and resolve interpretive issues.

Table 1: Sample Selection Procedures

Selection criteria	No. of Firms
1997 Fortune 500 firms named as PRP on SPIS database	245
Firms eliminated:	
Financial (SIC Division H)	(12)
Firms named to a site with no ROD issued	(27)
Firms not publicly traded during entire sample period	(11)
Firms added as PRP during the sample period	(13)
Final sample for analyses	182

The following procedures are performed for each year from 1991-1997:

1) a score of 1 is given for each disclosure item presented in the 10K (based on ED Index). Thus, the environmental disclosure score ranges from 0 (for no disclosure) to a maximum of 10 for years 1991 and 1992 (prior to SAB 92), and from 0 to 29 for years 1993-1997 (including disclosures required by SAB 92), and
2) the environmental disclosure score is divided by the total number of ED index items for each year (10 for years 1991 and 1992; 29 for years 1993-1997). The firms' final score for each sample year represents the percentage of GAAP disclosure (EDgaap). The variable (EDgaap) equally weights the disclosure items.

The independent variables are size, industry, profitability, and regulatory influence. Watts and Zimmerman (1978) suggest an association between company size and social responsibility disclosure. They argue that because political costs reduce management wealth, companies attempt to reduce costs by such devices as social responsibility disclosure campaigns. Since the magnitude of political costs is highly dependent on size, I hypothesize a positive relationship between size and ED (Watts and Zimmerman, 1978). The current study uses the natural log of total assets to proxy for size. Sensitivity analysis is conducted using two alternate size proxies (log of sales and log of market value of equity).

The current study uses dummy variables to identify firms in the five industries included in the Counsel on Economic Priorities (CEP) studies (petroleum, chemical, electric power, paper and pulp, and steel). The CEP identified these five industries as generating significant environmental hazards, thereby requiring substantial environmental disclosure. I therefore hypothesize that companies in these industries will disclose more than companies in other industries.

Table 2: Environmental Disclosure Index		
Item No.	Source	GAAP Disclosures
1	Reg. S-K Item 101	A general description of the business and specific disclosure of the effects that compliance with environmental laws, when material
2	Reg. S-KItem101	Estimated amount disclosed for capital expenditures representing current and succeeding fiscal years in which those expenditures may be material
3	Reg. S-KItem103	Disclosure of pending or contemplated administrative or judicial proceedings
4	Reg. S-KItem 303	Disclosure of material events and uncertainties known to management that would cause reported financial information to be unrepresentative of future operating results
5	SFAS5	Nature of accrual
6	SFAS5	Accrued amount
7	SFAS5	Nature of loss contingency
8	SFAS5	Estimate of additional possible loss or range
9	SFAS5	Statement that estimate cannot be made
10	SFAS5	Nature of probable unasserted claims that are possibly unfavorable
11	SAB92	Whether an asset is recorded for probable recovery
12	SAB92	Whether the accrual is undiscounted
13	SAB92	The discount rate used
14	SAB92	Expected payments for each of 5 succeeding years
15	SAB92	Reconciliation of the undiscounted to recognized amounts
16	SAB92	Material changes in expectations explained
17	SAB92	Circumstances affecting the reliability and precision of loss estimates
18	SAB92	Extent to which unasserted claims are reflected in any accrual or may affect the magnitude of the contingency
19	SAB92	Uncertainties with respect to joint and several liability
20	SAB92	Nature and terms of cost sharing arrangements with other PRPs
21	SAB92	Uncertainties with respect to insurance claims
22	SAB92	The extent to which disclosed but unrecognized contingent losses are expected to be recoverable through insurance, etc.
23	SAB92	Uncertainties about the legal sufficiency of insurance claims or solvency of insurance carriers
24	SAB92	The time frame over which accrued or unrecognized amounts may be paid out
25	SAB92	Material components of accruals and significant assumptions
26	SAB92	Recurring costs associated with managing hazardous substances and pollution in ongoing operations
27	SAB92	Mandated expenditures to remediate previously contaminated sites
28	SAB92	Other infrequent or nonrecurring cleanup expenditures, anticipated but not required in the present circumstances
29	SAB92	Loss disclosure with respect to particular environmental sites that are individually material

Cowen et al. (1987) cite profitability as a factor that allows or impels management to provide more extensive social responsibility information. This argument may be related to Zmijewski and Hagerman's (1981) contention that higher net income increases firm visibility. In addition, SEC enforcement activities related to environmental liabilities tend to concentrate on the "deep pocketed" or more profitable firms. Therefore, I hypothesize a positive relationship between profitability and ED. The current study uses return on assets to proxy for profitability.

Similar to Barth et al. (1997), an environmental liability estimate based on data provided by the EPA proxies for regulatory influence. The estimated liability identified by the EPA captures the effect of regulatory pressure or rules pertaining to environmental disclosure. I hypothesize larger liabilities to be associated with increased disclosure. The environmental liability is based on information provided in the ROD, which provides estimated costs of cleanup for Superfund sites. Barth et al. (1997) intend to capture the combined effect of regulatory pressure or rules pertaining to disclosure in general and to environmental disclosure in particular. The following procedures are applied:

1. The Superfund PRP Listing is used to identify the number of sites to which each sample firm is named as of 12/31/1997.
2. The SPIS database is used to obtain the RODs for all sites to which sample firms are named as of 1997. The ROD is examined for each site to obtain the Present Worth Cost (PWC), which represents the present value of the estimated clean up costs for the site.
3. For each site, the total number of PRPs is determined by sorting the Superfund PRP Listing by site number.
4. The number of publicly traded PRPs for each site is obtained using the EDGAR database, and is used to compute an average liability for each site. The average liability is calculated by dividing the PWC by the number of publicly traded PRPs for each site, which indicates the potential for shared responsibility for cleanup.
5. The average liability for each site to which a sample firm is named is then added to obtain a total average liability. The current study uses total average liability as a proxy for potential environmental liability.

The current study estimates the model using ordinary least squares regression (OLS), which requires normally distributed error terms. To increase the probability of normality in the error terms, the variables are transformed using the natural logarithm to allow the distribution to approach normality. The final form of the variables in the base model includes log of total assets (SIZE), log of total average liability, return on assets (ROA) and dummy industry variables. Descriptive statistics for model variables are summarized at Table 3.

Table 3: Descriptive Statistics for Model Variables

Variable	N	Minimum	Mean	Median	Maximum	St. Dev
EDgaap (ratio)	1274	0.000	0.356	0.345	0.900	0.238
Disclosure score	1274	0.000	7.990	8.000	18.000	5.180
Total assets (millions)	1274	119	14,465	6,201	304,012	30,269
Log of total assets	1274	4.780	8.886	8.732	12.620	1.037
ROA (ratio)	1274	-0.619	0.043	0.043	0.600	0.071
Liability (millions)	1274	0.026	17.532	6.099	243.929	29.438
Log Liability	1274	-3.660	1.729	1.808	5.500	1.705
Sales (millions)	1274	19	12,431	6,471	178,174	19,387
Log of sales	1274	2.940	8.899	8.775	12.090	0.945
MVE (millions)	1209	7	12,176	5,004	239,539	20,309
Log MVE	1209	1.970	8.551	8.518	12.390	1.406

The dependent variable (EDgaap) ranges from 0% to 90%, with a mean of 36%. The varying frequencies reveal that firms exercised much discretion in their disclosure practices. Disclosure scores range from 0 to 18, with a mean of 8. Log of assets ranges from 4.78 to 12.62 (in dollars, $119 million to $304 billion), which indicates that the sample consists of large firms. ROA ranges from -.619 to .600, indicating that some sample firms were not profitable, which may affect the ability to pay. Log of total average liability ranges from –3.660 to 5.5 (in dollars, $26,000 to $244 million), which highlights the magnitude of potential environmental obligations after being adjusted to reflect potential shared responsibility.

For sensitivity analysis, the current study uses two alternate proxies for size (log of sales and log of market value of equity). The log of sales ranges from 2.94 to 12.09 (in dollars, $19 million to $178 billion), and log of market value of equity ranges from 1.97 to 12.39 (in dollars, $7 million to $240 billion).

The hypotheses in the alternate form are:

H1: There is a positive association between company size and the amount of ED.

H2: Companies in the petroleum, chemical, electric power, paper and pulp, and steel industries provide more ED than companies in other industries.

H3: There is a positive association between profitability and the amount of ED.

H4: There is a positive association between regulatory influence and the amount of ED.

The complete specification of the regression is EDgaap equals alpha 0 plus alpha 1 SIZE plus alpha 2 CHEM plus alpha 3 OIL plus alpha 4 PAPER plus alpha 5 STEEL plus alpha 6 POWER plus alpha 7 ROA plus alpha 8 total average liability plus an error term, where, Size equals log of total assets; Industry equals indicators that equal 1 for classifications in oil (2-digit SIC 13), chemical (2-digit SIC 28), power (2-digit SIC 49), paper and pulp (2-digit SIC 26), and steel industries (2-digit SIC 33), and 0 otherwise; Profitability equals return on assets (ROA); Regulatory Influence equals total average liability (environmental liability proxy based on liability estimates provided in EPA Records of Decision).

RESULTS

Table 4 provides the estimation results for the current study. The p-values for all model coefficients are based on White's (1980) heteroskedasticity-corrected standard errors. The adjusted R squared for the base model (Model 1) is .181 (p-value .000). The results for Model 2, which introduces industry indicators, show a larger adjusted R squared (.360, p-value .000) than the base model. The comparison of the two models illustrates the incremental value of the industry indicators. The adjusted R squared is improved when industry dummy variables are included in model (Model 2).

The reported findings are insensitive to outliers as identified by the Belsley, Kuh, and Welsch (1980) DFFITS statistics. No DFFITS statistic exceeds the size-adjusted cutoff; therefore, outlier analysis reveals no influential data affecting model coefficients. The Durbin Watson statistics for Model 1 (1.95) and Model 2 (2.01) do not indicate autocorrelation (Studenmund, 1997). Condition Numbers below 30 for Model 1 (22.34) and Model 2 (24.31) do not indicate severe multicollinearity (Studenmund, 1997). As all test statistics are reported using White (1980) standard error estimates, standard diagnostics do not reveal significant problems with the model specifications.

The primary results (Model 2) indicate a positive association between company size and the amount of ED (p-value =.000). The findings support the current hypothesis (H1), and are consistent with prior research (Barth et al., 1997; Patten, 1992, 1991; Cowen et al., 1987). The findings indicate that firms attempt to reduce political costs by increasing social disclosure. Since the magnitude of political costs is highly dependent on size, Watts and Zimmerman (1978) hypothesized that there is positive relationship between size and social disclosure.

The results indicate a negative association between profitability and the amount of ED (p-value = .000). Prior studies that examine the impact of profitability on social disclosure find no association; however, these studies do not specifically include ED (Cowen et al., 1987; Patten,

1990). Since, the "deep pocketed" party is often legally liable for all remediation costs under the Superfund Act, the current study posits that profitability would impel management to disclose more about the environmental activities of the firm (H3). However, the negative association found between profitability and the amount of ED suggests that less profitable companies disclose more. This finding may be explained within the context of voluntary disclosure literature, which focuses on management's earnings forecasts (Pownall et al., 1993; Skinner, 1994, 1995).

Table 4: Regression Results

EDgaap	Pred.Sign	Model(1)	Pred.Sign	Model(2)
N		1274		1274
Intercept		0.101(0.069)		0.076(0.137)
SIZE	+	0.021(0.000)	+	0.017(0.003)
ROA	+	-0.468(0.000)	+	-0.524(0.001)
LIAB	+	0.051(0.000)	+	0.044(0.001)
CHEM			+	0.186(0.000)
OIL			+	0.225(0.001)
PAPER			+	0.0984(0.001)
STEEL			+	0.188(0.001)
POWER			+	0.289(0.001)
Adj. R^2		0.181		0.360
Prob F		0.000		0.000

P-value using White's standard errors. Signed tests are one-tailed.

Legend:

SIZE	Log of total assets
ROA	Net income/book value of total assets
LIAB	Log of total average liability based on EPA estimates
CHEM	1 if firm is in chemical industry (2-digit SIC 28), 0 otherwise
OIL	1 if firm is in oil industry (2-digit SIC 13), 0 otherwise
PAPER	1 if firm is in paper industry (2-digit SIC 26), 0 otherwise
STEEL	1 if firm is in steel industry (2-digit SIC 33), 0 otherwise
POWER	1 if firm is in power industry (2-digit SIC 49), 0 otherwise

Skinner (1994, 1995) provides evidence consistent with the idea that firms with negative earnings news tend to voluntarily disclose bad news. Pownall, Wasley, and Waymire (1993) also provide evidence that voluntary earnings disclosures tend to convey bad news. Since much variation exists in the ED practices of firms, ED may be explained in the context of voluntary disclosure. Environmental liability exposure may be viewed as bad news, and less profitable firms may be more

likely to disclose ED. This is also consistent with more recent studies that suggest there are legal incentives for firms to disclose bad news (Skinner, 1997; Evans et al., 2002).

The results reveal a positive association between the log of total average liability and the amount of ED, which indicates that firms with higher environmental liabilities disclose more (p-value = .000). Since the total average liability measure is based on EPA liability estimates, this finding reflect firms responding to regulatory influence, and supports the current hypothesis (H4). The results indicate that companies in the petroleum, chemical, paper and pulp, and steel industries provide more ED than companies in other industries (p-values =.000). The findings support the current hypothesis (H2), and are consistent with prior studies that suggest that companies in certain environmentally sensitive industries have greater incentive for projecting a positive social image, and therefore disclose more (Bowman and Haire, 1976; Heinze, 1976; Cowen et al., 1987; Patten, 1991).

The current study uses additional size proxies to check the robustness of the primary results. Separate regressions are run for the base model (Model 1), substituting log of total assets with log of sales (Model 1a) and log of market value of equity (Model 1b). The results indicate that the log of sales (p-value =.083) and log of market value of equity (p-value = .057) are not associated with the percentage of environmental disclosure. The results are contrary to the expectations of the current study, and are not consistent with the primary results. Thus, the primary results are not robust to alternate specifications of size. While log of total assets is associated with environmental disclosure, log of sales and log of market value of equity are not.

Empirical results indicate that the level of environmental disclosures is associated with size, profitability, industry, and regulatory influence. As firms continue to exercise much discretion with respect to environmental disclosure, understanding the determinants of environmental disclosure choice may help regulators in their monitoring and enforcement activities. In addition, the findings may help standard setters determine whether more guidance is needed specifically related to environmental liabilities. Socially conscious investors may find this information useful when evaluating firms' environmental disclosures.

DISCUSSION AND CONCLUSIONS

The current study contributes to the understanding of ED practices of publicly traded U.S. firms because it examines the extent of GAAP disclosures in 10K filings, as well as factors associated with the level of ED. Using a sample of firms that have been identified as potentially responsible parties by the EPA, the current study finds that the extent of environmental disclosure varies between sample firms. Although disclosure scores increased overall during the sample period, the percentage of GAAP disclosures decreased significantly. Firms continued to exercise much discretion. Knowledge of such discretion is important to investors, creditors, regulators and standard setters that expect complete and consistent environmental disclosure practices of public companies.

Policy implications for the SEC include improving efforts to promote adequate environmental disclosures. For the FASB and AICPA, the findings indicate that current authoritative guidance is not effective in promoting consistent and adequate environmental disclosure. Therefore, standard-setters must endeavor to provide comprehensive and specific guidance regarding environmental liabilities that can be consistently applied by firms and readily monitored by the SEC through use of EPA data.

The findings also have implications for creditors and investors, who expect transparency in the disclosures of public companies. The transparency of disclosure in 10K reports is now particularly important in light of the increased attention on financial reporting due to the current wave of corporate secrecy and corruption. The findings of the current study may alert creditors and investors to the inconsistencies in the environmental disclosures of publicly traded firms, further contributing to public distrust of financial reports and the accounting profession.

The current study examines factors related to environmental liability disclosure decisions by estimating the association between ED and proxies for size, industry, profitability, growth and regulatory influence. The findings indicate that size, industry, profitability, and regulatory influence are significantly associated with environmental disclosure decisions. Understanding which firm characteristics are associated with the level of disclosure can assist regulators in better monitoring the disclosure practices of firms. In addition, such understanding can assist investors when evaluating the adequacy of environmental disclosure in financial statements. Researchers interested in building a theory of social disclosure can also benefit from the current findings.

The study contributes to the corporate social reporting literature by providing more recent information, specifically in the context of environmental disclosure. Early literature that examines the determinants of social disclosure choice did not focus exclusively on environmental disclosure. Whereas early studies found no association between profitability and the extent of social disclosure, the current study finds a statistically significant negative association between profitability and environmental disclosure. The study examines a longer period and finds profitability, growth, size, industry classification and environmental liability estimates to be associated with the level of environmental disclosure.

ACKNOWLEDGMENT

Thanks always to my parents Annie and Andrew Cox, for their unending support. I want to also thank each member of my Dissertation Committee at Virginia Commonwealth University: Dr. Ruth W. Epps (Chair), Dr. Benjamin B. Bae, Dr. Kenneth N. Daniels, Dr. Tanya S. Nowlin, and Dr. Benson Wier, for their guidance and direction. In addition, I would like to express my deepest gratitude to the faculty at George Mason University, who provided much needed assistance. Finally, thanks to the faculty in the AAA Mid-Atlantic Region, for providing excellent feedback.

REFERENCES

American Institute of Certified Public Accountants (AICPA), (1996). *Statement of Position 96-1: Environmental Remediation Liabilities* (Including Audit Guidance). New York, NY.

Bank, D. (Aug 22, 2002). Group urges enforcing rules of environmental disclosure, *The Wall Street Journal*, B2.

Barth, M. E., McNichols, M. F. & Wilson, G. P. (1997). Factors associated with firms' disclosures about environmental liabilities. *Review of Accounting Studies* 2: 35-64.

Beams, F. A. & Fertig, E. (1971). Pollution control through social cost conversion. *Journal of Accounting* 132(5): 37-42.

Belsley, D.A., Kuh, E. & Welsch, R. E. (1980). *Regression Diagnostics.* New York: John Wiley and Sons.

Berthelot, S., Cormier D. & Magnan M.. (2003). Environmental Disclosure Research: Review and Synthesis. *Journal of Accounting Literature* 22: 1-44.

Bibler. G.A. & Davis, C.P. (2003). Disclosing environmental liabilities in the wake of Sarbanes-Oxley, *The Metropolitan Corporate Counsel*, (4) 1-4.

Bowman, E. H. & Haire, M. (1975). A strategic posture toward corporate social responsibility. *California Management Review* 18(2): 49.

Churchill, T. A. (1978). Reporting on corporate social responsibility: A progress report. *Journal of Contemporary Business* 7(1): 5-18.

Committee on Commerce, (1995). *Reform of Superfund Act of 1995: Hearings Before the Subcommittee on Commerce, Trade and Hazardous Materials of the Committee on Energy and Commerce.* Washington DC, U.S. Government Printing Office.

Cowen, S., Ferreri, L. & Parker, L. (1987). The impact of corporate characteristics on social responsibility disclosure: A typology and frequency-based analysis. *Accounting, Organizations and Society* 12(2): 111-122.

Cox, R. M. Jr. (2001). Tough environmental regulations bring new opportunities for CPAs. Pennsylvania *CPA Journal* 72(2): 16-17.

Estes, P. (1976). Standards for corporate social reporting. *Management Accounting* 58(5): 19-22.

Evans III, J. H. & Sridhar, S. S. (2002). Disclosure-disciplining mechanisms: Capital markets, product markets, and shareholder litigation. *The Accounting Review* 77(3): 595-626.

Financial Accounting Standards Board. (1975). *Accounting for Contingencies, FASB Statement No. 5*, Stamford, CT: FASB.

Fitzgerald, P. (1996). Big investment chill. *Chemical Marketing Reporter* 249(22): SR6.

Freedman, M. & Wasley, C. (1990). The association between environmental performance and environmental disclosure in annual reports and 10Ks. *Advances in Public Interest Accounting*: 183-193.

Gamble, G. O., Hsu, K. & Radke, R. (1995). Environmental disclosures in annual reports and 10Ks: An examination. *Accounting Horizons* 9(3): 34-54.

Heinze, D.C. (1976). Financial correlates of a social involvement measure. *Akron Business and Economic Review* 7(1): 48-51

Herdman, R.K. (2002). *Testimony Concerning the Roles of the SEC and the FASB in Establishing GAAP Before the House Subcommittee on Capital Markets, Insurance, and Government Sponsored Enterprises.* Committee on Financial Services. (May 14).

Ingram, R. & Frazier, K. (1980). Environmental performance and corporate disclosures. *Journal of Accounting Research* 18(2): 614-622.

Johnson, L. T. (1993). Research in environmental reporting. *Accounting Horizons* 7(3): 118-123.

Kreuze, J. G., Newell, G. E. & Newell, S. J. (1996). What companies are reporting. *Management Accounting* 78(1): 37-43.

Lavelle, M. (1992). Superfund studies begin to fill hole in data-dry field. *Business Watch, The National Law Journal* 14(20): 19-20.

Nader, R. (1978). Corporate disclosure: The public right to know. *Journal of Contemporary Business* 7(1): 25-32.

Patten, D.M. (1992). Intra-industry environmental disclosures in response to the Alaskan oil spill: A note on legitimacy theory. *Accounting, Organizations and Society* 17(5): 471-475.

Patten, D.M. (1991). Exposure, legitimacy, and social disclosure. *Journal of Accounting and Public Policy* 10(4): 297-308.

Patten, D.M. (1990). The market reaction to social responsibility disclosures: The case of the Sullivan Principles signings. *Accounting Organizations and Society* 15(6): 575-587.

Pownall, G., Wasley, C. & Waymire, G. (1993). The stock price effects of alternative types of management earnings forecasts. *Accounting Review* 68(4): 896-912.

Price Waterhouse, (1994). *Progress on the environmental challenge: A survey of corporate America's environmental accounting and management*. New York, Price Waterhouse.

Price Waterhouse, (1992). *Accounting for environmental compliance: Crossroad of GAAP, Engineering, and Government.* New York, Price Waterhouse.

Probst, K. N., Fullerton, D., Litan, R. E. & Portney, P. R. (1995). *Footing the bill for superfund cleanups: Who pays and how?* Washington DC: The Brookings Institution and Resources for the Future.

Rockness, J. W. (1985). An assessment of the relationship between US corporate environmental performance and disclosure. *Journal of Business, Finance and Accounting* 12(3): 339-354.

Sack, R. J., Boatsman, J. R., Fell, R. S., Krogstad, J. L., Martin S. J. & Niles, M. S. (1995). American Accounting Association's Securities and Exchange Commission Liaison Committee: Mountaintop issues: From the perspective of the SEC. *Accounting Horizons* 9(1): 79-87.

Securities and Exchange Commission, (1993). Staff Accounting Bulletin No. 92. *Federal Register* (June 14, 1993).

Securities and Exchange Commission, (2000). Regulation S-K. *Federal Register* (April 11, 2000)

Sissell, K. (Nov 13, 2002). Senators query SEC on policy for environmental disclosures, *Chemical Week* 164 (45), 14.

Skinner, D. J. (1994). Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32(1): 38-60.

Skinner, D.J. (1995). Do the SEC's safe harbor provisions encourage forward-looking disclosures? *Financial Analyst Journal* 51(4): 38-44.

Skinner, D.J. (1997). Earnings disclosure and stockholder lawsuits. *Journal of Accounting and Economics* 23(3)): 249-282.

Spicer, B. (1978). Investors, corporate social performance and information disclosure: An empirical study. *The Accounting Review* 53(1)): 94-111.

Stanny, E. (1998). Effect of regulation on changes in disclosure of reserved amounts for environmental liability. *Journal of Financial Statement Analysis* 3(4): 34-49.

Studenmund, A.H. (1997). *Using econometrics: A practical guide* (3rd Edition). Reading, Massachusetts: Addison-Wesley.

Trotman, K. T. & Bradley, G. W. (1981). Associations between social responsibility disclosure and characteristics of companies. *Accounting, Organizations and Society* 6(4): 355-362.

Walden, W. D. & Schwartz, B. N. (1997). Environmental disclosures and public policy pressure. *Journal of Accounting and Public Policy* 16(2): 125-154.

Watts, R. L. & Zimmerman, J. L. (1978). Towards a positive theory of the determination of accounting standards. *The Accounting Review* 58(1): 112-134.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48(4): 817-838.

Wiseman, J. (1982). An evaluation of environmental disclosures made in corporate annual reports. *Accounting, Organizations and Society* 7(1): 53-63.

Zmijewski, M. E. & Hagerman, R. L. (1981). An income strategy approach to the positive theory of accounting standard setting/choice. *Journal of Accounting & Economic* 3(2): 129-150.

THE RELATION BETWEEN PURCHASED GOODWILL AND TARGET CHARACTERISTICS

Natalie Tatiana Churyk, Northern Illinois University
Katrina L. Mantzke, Northern Illinois University

ABSTRACT

This study investigates the relation between the goodwill recorded in a purchase transaction and the characteristics of the target firms that are acquired. While prior research on goodwill has largely ignored the characteristics of the target firm, this study contributes to the literature by explicitly investigating the relation between the goodwill booked by the acquirer and the target characteristics that may explain why an acquirer paid a premium for the target. Our results provide evidence that acquirers pay a premium for the targets' statutory-based characteristics (intangible assets other than goodwill) and the targets' corporate organizational and financial characteristics (tax loss carryforwards, lower borrowing rate, and lower debt-to-equity ratio). Our results provide only weak evidence that goodwill is positively related to technology-based characteristics and workforce-based characteristics of the targets. Finally, our results fail to show that goodwill is positively related to the customer/market-based characteristics of the target.

INTRODUCTION

Generally accepted accounting principles (GAAP) require acquiring firms to record goodwill as an asset when they purchase a target firm and pay more than the fair market value of the identifiable net assets of that firm. Researchers have studied goodwill from a number of different perspectives. Early researchers debated the proper accounting for and reporting of goodwill (Emery 1951; Miller, 1973; Colley and Volkan, 1988; Ma and Hopkins, 1988). Other researchers were concerned with what goodwill represents. Some have interpreted goodwill to be the value of expected excess future earnings discounted over a number of years (Walker, 1938; Emery, 1951; Nelson, 1953; Gynther, 1969; Ma and Hopkins, 1988). Another stream of research interprets goodwill to be a momentum or an initial push comprised of unrecorded aspects of the firm, such as managerial skill, economies of scale, and customer relations (Emery, 1951; Nelson, 1953; Barlev, 1973). The empirical studies that followed considered goodwill from the perspective of the acquirer and found goodwill to be value relevant (Chauvin and Hirschey, 1994; McCarthy and Schneider, 1995; Jennings, Robinson, Thompson, and Duvall, 1996; and Vincent, 1997).

Yet, the foregoing research is lacking. No one has studied goodwill to investigate the relation between the goodwill booked by the acquirer and the target characteristics that may explain why an

acquirer paid a premium for the target. This paper studies this relation to provide a more complete picture of why purchase premiums exist.

Statement of Financial Accounting Standards (SFAS) No. 141 (2001) specifies the generally accepted accounting principles for business acquisitions. This standard states that acquirers must record intangible assets that arise from "contractual or other legal rights" or those that are "capable of being separated or divided from the acquired entity" (par. 39). Intangible assets not meeting these recognition criteria must be recorded as goodwill. This catchall asset is defined in paragraph 43 of SFAS No.141 as "the excess of the cost of an acquired entity over the net of the amounts assigned to assets acquired and liabilities assumed."

To further investigate what goodwill is comprised of, we empirically examine the relation between goodwill and the *target* characteristics that motivate acquisitions. We have grouped these characteristics into five categories that are described as follows. 1) Customer/market-based characteristics are measured in terms of vertical integration, horizontal integration, and geographic areas of major lines of business. 2) Technology-based characteristics are measured by research and development expenditures. 3) Statutory-based characteristics are measured by intangibles other than goodwill. 4) Work-force based characteristics are measured in terms of managerial talent. 5) Corporate organizational/financial characteristics are measured by cash reserves, tax loss carryforwards, borrowing rates, and debt-to-equity ratios.

This study extends the literature by empirically examining *target* characteristics underlying goodwill recognition, as opposed to the characteristics of the acquirer. Our results indicate that both statutory-based characteristics and corporate organizational/financial characteristics are significantly positively related to goodwill when all characteristics are included in the model. Other characteristics become significantly positively related to goodwill when subsets of the categories are examined due to data availability. In providing evidence of these relations, this research provides new insight as to which target characteristics compel acquiring firms to pay a premium.

The paper is organized as follows. In the next section, we discuss the characteristics of target firms that prior research suggests are related to goodwill and develop hypotheses. The third section describes the sample and the cross-sectional model relating goodwill to the acquired-firm characteristics. A discussion of the results and conclusions follows.

LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

The notion that goodwill is comprised of firm characteristics has had a slow evolution. Nelson (1953) is the first to describe goodwill as a composite of characteristics, the value of which dissipates over time. As required by GAAP, firms expense some of these items as they are incurred, e.g., advertising costs, capitalize others as intangible assets, e.g., trademarks, and do not recognize others because of measurement difficulties, e.g., customer lists.[1] Thus the economic benefit associated with many firm characteristics is not recognized in financial statements unless the firm is purchased as a going concern. At that time, these benefits may be reflected in recorded goodwill.

To complement the view that goodwill is a composite of firm characteristics, Barlev (1973) proposes that mergers provide economic gains (i.e., synergy) justifying a premium over market value. The economic gains are linked with type of merger: horizontal, vertical, or conglomerate. Numerous studies support the synergy motives for mergers and acquisitions, including Falk and Gordon (1977), Jensen and Ruback (1983), Weston and Halpern (1983), Bradley, Desai, and Kim (1988), Berkovitch and Narayanan (1993), Maquieira, Megginson, and Nail (1998), and Zhang (1998).

Tax benefits have also been linked to goodwill. Haw, Pastena, and Lilien (1987) find that higher premiums are paid for firms with tax loss carryforwards than for similar firms without carryforwards. Hayn (1989) finds that higher premiums compensate target firm's shareholders who are faced with an immediate capital gains tax. Further evidence supports the hypothesis that the tax shield resulting from increased leverage is a merger benefit (Lewellen, 1971; Lintner, 1971; Melnik and Pollatschek, 1973; Scott, 1977; Amihud and Lev, 1981; Walkling and Edmister, 1985; Crawford and Lechner, 1996; Raad, Ryan, and Sinkey, 1999).

Finally, empirical support for the Nelson (1953) view of goodwill's association with certain characteristics is provided by Chauvin and Hirschey (1994). They find that goodwill is value relevant. They document a positive association between goodwill and *post-acquisition* advertising and research and development in all model specifications. They also find a positive association between goodwill and *post-acquisition* intangible assets in some model specifications. Yet, these results do little to further our understanding of which *target* characteristics compel acquirers to pay premiums for target firms.

This prior research lays the foundation for categorizing target characteristics in order to measure the relation between these characteristics and the goodwill recorded in the acquisition. Employing the categories used in SFAS No. 141 and its predecessor Exposure Draft, intangibles are grouped into the following categories: customer/market-based characteristics, technology-based characteristics, statutory-based characteristics, workforce-based characteristics, and corporate organizational and financial characteristics. GAAP requires separately identifiable intangibles other than goodwill to be amortized over their stipulated lives. If any particular intangible cannot be measured with sufficient reliability, it is included in goodwill. In addition, other amounts are also included in goodwill because it is the most expedient and least costly accounting treatment.

Customer/market-based Characteristics

Characteristics in this category comprise synergies that result from the business combination as they relate to delivery systems and distribution channels, advertising, and existing presence in desirable geographic locations or markets. Economies of scale in production, advertising, and distribution arise from horizontal integration. Both horizontal and conglomerate mergers allow for expansion into new markets or new territories. Economies of scope produce cost savings in vertical integration. All three types of mergers can result in increases in market or monopoly power. Despite the benefits expected from such combinations, the items involved either are non-identifiable

intangibles or are expensed as incurred under GAAP. Yet, prior research demonstrates that these items do have value. For example, Doukas and Travlos (1988) find that shareholders of multinational firms not already operating in the country of the target firm benefit from expansion into the target's industry or geographic market. In a similar vein, Hirschey and Weygandt (1985) provide empirical evidence that the benefits of current advertising outlays carry over to subsequent periods.

Customer/market-based characteristics, i.e., synergy, expansion into new territories, and advertising, are expected to result in benefits to the purchasing firm and are positively associated with goodwill. Stated in the alternative form, our first hypothesis is as follows:

H1: Customer/market-based characteristics of the target firm, measured on the most recent annual report date prior to the purchase, are positively related to goodwill measured on the purchase date.

Support for this hypothesis would provide evidence that acquirers are willing to pay a premium for the customer/market-based characteristics of target firms.

Technology-based Characteristics

Technology-based characteristics include items such as computer software and licenses, computer programs, information systems, secret formulas, and research and development (R&D). Nelson (1953) predicts that such items are associated with goodwill. All types of mergers (vertical, horizontal, and conglomerate) enable access to proprietary technology for the acquirer. Although many of the characteristics themselves are difficult to measure, the amount spent on R&D (which gives rise to such assets) is often provided in the annual report.

Prior research demonstrates that R&D expenditures from the current period do benefit future periods (see Hirschey and Weygandt (1985), for instance). However, because GAAP requires the immediate write-off of R&D, the benefits arising from R&D activities are likely understated. To the extent the acquirer values these aforementioned benefits, they will be reflected in the purchase price. Therefore, we predict that technology-based characteristics of the target (as measured by R&D) are positively associated with goodwill. Stated in the alternative form, the second hypothesis is as follows:

H2: Technology-based characteristics of the target firm, measured on the most recent annual report date prior to the purchase, are positively associated with goodwill measured on the purchase date.

Support for this hypothesis would provide evidence that acquirers are willing to pay a premium for technology-based characteristics of target firms.

Statutory-based Characteristics

Statutory-based characteristics include patents, copyrights, franchises, trademarks, and trade names. These items constitute intangible assets other than goodwill. However, these assets are only recorded as such when they are purchased; they are not booked as assets when developed internally. To the extent a target's statutory-based characteristics are valuable in the acquisition of a going concern and are either undervalued or are not recorded by the target firm, they will appear in goodwill of the combined firm. Nelson (1953) predicts that these items do generate goodwill. Chauvin and Hirschey (1994) predict that such intangibles prolong the life of goodwill but fail to find a significant relation between goodwill and intangible assets of the acquirer. We investigate the relation between statutory-based characteristics of the target and goodwill. Stated in the alternative form, the third hypothesis is as follows:

H3: Statutory-based characteristics of the target, measured on the most recent annual report date prior to the purchase, are positively associated with goodwill measured on the purchase date.

Support for this hypothesis would provide evidence that acquirers are willing to pay a premium for statutory-based characteristics of target firms.

Workforce-based Characteristics

Workforce-based characteristics may include an assembled and trained staff, non-union status, strong labor relations, favorable wage rates, superior management or other key employees, technical expertise, and ongoing training and recruiting programs. Based on Nelson's (1953) work, Chauvin and Hirschey (1994) argue that managerial scope economies accrue to organizations and result in economic goodwill. However, they fail to find a relation between the market share of the acquirer, a proxy for managerial talent, and goodwill. We investigate this relation further by testing for a positive relation between goodwill and workforce-based characteristics (as measured by firm sales compared to industry sales). Stated in alternative form, the fourth hypothesis is as follows:

H4: Workforce-based characteristics of the target, measured on the most recent annual report date prior to the purchase, are positively associated with goodwill measured on the purchase date.

Support for this hypothesis would provide evidence that acquirers are willing to pay a premium for workforce-based characteristics of target firms.

Corporate organizational/financial Characteristics

Corporate organizational/financial characteristics include the savings value of escrow funds, favorable financial arrangements, outstanding credit ratings, fundraising capabilities, access to capital markets, and favorable government relations. Jensen (1986) proposes that free cash flow (cash more than the amount necessary to fund positive net present value projects) imposes agency costs. To maximize shareholder wealth, managers must either invest excess cash in positive net present value projects or return the cash to shareholders in the form of dividends. However, because managers tend to squander cash, an acquirer may be able to employ more efficiently the excess financial characteristics of the target. Smith and Kim (1994) find that the market views acquisitions of targets with free cash flow more positively than targets without free cash flow. However, it is possible that the acquirer purchases the target based on the mistaken belief that it can better manage the target, when in fact this is not true. This situation has been referred to as hubris (Roll, 1986). Although Falk and Gordon (1977) find a correlation between large target cash reserves and goodwill, hubris may result in a lack of correlation.

Credit ratings are heavily influenced by the issuer's capacity to repay its long-term financial obligations. A firm with superior credit has an easier time issuing debt or securing loans than a firm with poor credit. Furthermore, business combinations result in the ability of the new entity to raise funds at a lower cost because the risk associated with larger firms is typically lower than that associated with smaller firms. This may be particularly important if the target is less risky than the parent and has a better credit rating and lower borrowing rates. We predict a positive relation between the target's credit rating and goodwill and between the parent-to-target borrowing rate ratio and goodwill.

Haw et al. (1987) predict that higher merger premiums result for firms with tax loss carryforwards than for firms without. Also, the tax benefits of debt capacity may be reflected in goodwill. Evidence supports the theory that the existence of leverage-related tax shields is associated with larger merger premiums (Lewellen, 1971; Lintner, 1971; Melnik and Pollatschek, 1971; Scott, 1977; Amihud and Lev, 1981; Walkling and Edmister, 1985; Raad, Ryan, and Sinkey, 1999). We predict a positive relation between goodwill and tax-related financial characteristics of the target. In summary, we expect a positive relation between corporate-organizational/financial characteristics and goodwill. Stated in the alternative form, the fifth hypothesis is as follows:

H5: Corporate organizational/financial characteristics of the target, measured on the most recent annual report date prior to the purchase, are positively associated with goodwill measured on the purchase date.

Support for this hypothesis would provide evidence that acquirers are willing to pay a premium for corporate organizational/financial characteristics of target firms.

METHODOLOGY

Sample Selection

Table 1 summarizes the process used to collect our sample. The 1,555 firms delisted from Compustat due to mergers/acquisitions for the period 1996-1998 were identified as target firms for the study. Firms were eliminated from this initial sample for the following reasons: (1) insurance companies and financial institutions, i.e., firms with SIC codes between 6000 and 6999 inclusive (496 targets and 115 parents), due to specialized accounting requirements and/or regulation; (2) 78 firms whose parent companies could not be identified; (3) 225 firms due to the use of the pooling-of-interests accounting method; (4) 465 firms that had missing acquisition details; (5) one firm that was both a target and a parent in separate years[2]; and (6) nine firms for which the necessary financial statements could not be located. These eliminations resulted in a sample of 166 potential firms with purchased goodwill details. To minimize the effect of outliers, we followed Amir, Kirshenhieter and Willard (1997) and eliminated 17 observations (9.7%) which had an absolute value of the R-student statistic (regression residual divided by the residuals' standard error) greater than 3.0 in all analyses. The final sample includes 149 firms. Because data are not available to proxy each characteristic of each target firm included in this study, individual regressions use samples ranging in size from 51 to 148 observations.

Table 1: Sample Selection

	1996	1997	1998	Total
Delisted firms	446	517	592	1,555
Target SIC = 6000	(165)	(165)	(166)	(496)
Cannot identify parent	(26)	(16)	(36)	(78)
Pooling method used	(64)	(69)	(92)	(225)
Parent SIC = 6000	(4)	(3)	(108)	(115)
Firm is target and parent	(1)	--	--	(1)
Missing goodwill details	(140)	(205)	(120)	(465)
Missing parent data	(4)	(1)	--	(5)
Missing target data	(2)	(2)	--	(4)
Outlier Removal	(3)	(6)	(8)	(17)
Characteristics Sample	37	50	62	149

Descriptive statistics for our sample of target firms and for the population of target firms (other than firms with SIC codes between 6000 and 6999) during the same time period (1996-1998) are presented in Table 2.

Table 2: Descriptive Statistics – Target Sample vs. Population of Delisted Firms Due to Merger/Acquisition Activity (in millions)[a] and Industry Concentration

Panel A:

	Sample[a]		Population[b]		Comparison[c]
Measure	Mean	Median	Mean	Median	Prob>|t|
Total assets	508.47	131.67	629.36	118.47	0.5916
Market value	487.35	88.47	565.93	112.07	0.6367
Sales	535.94	146.83	538.60	119.44	0.9890
Debt to Equity	130.56%	49.04%	87.67%	37.95%	0.1831
Return on Assets	-1.09%	2.96%	-4.55%	2.61%	0.1809
Return on Equity	-5.04%	7.11%	2.39%	6.88%	0.5970

Panel B:

SIC Code	SIC Description	Sample/Frequency %		Population/Frequency %	
<1000	Agriculture, forestry, and fishing	1	1%	8	1%
1000-1999	Mining and construction	7	5%	74	8%
2000-2999	Manufacturing – food, tobacco, textile, apparel, lumber, furniture, paper, printing, chemicals and refining.	21	14%	118	12%
3000-3999	Manufacturing – rubber, leather, stone, metal, machinery, electronic, trans-portation, controlling instruments, miscellaneous	47	32%	268	28%
4000-4999	Transportation, communications, electric, gas and sanitary	17	11%	121	13%
5000-5999	Retail trade	19	13%	110	11%
7000-7999	Services – hotels, personal, business automotive repair, motion picture, amusement	23	15%	198	20%
8000-8999	Services – health, legal, educational, social, museums, membership, accounting, engineering, research	14	9%	70	7%
Total		149	100%	967	100%

[a] The sample consists of 149 observations obtained from the period 1996 through 1998. Not all data items are available for every firm in the sample. Observations vary from 137 to 149 per three-year period.

[b] The population was derived as follows: All delisted firms reported in the Research files of Compustat – 5197. Removed 1149 firms with SIC codes between 6000 and 6999. Removed 2989 firms for years other than the years studied in this paper (1996-1998). Removed 92 firms for which no data was available on Compustat leaving a total of 967 potential population firms. Not all data items are available for every firm in the populations. Observations vary from 945 to 967 per three-year period.

[c] The comparison column contains results of a t-test comparing the mean of the sample to the mean of the populations to determine if the two groups are statistically equivalent.

The sample does not differ from the population in terms of size (assets, market value of equity, sales), leverage (debt-to-equity), or profitability (return-on-assets, return-on-equity). Furthermore, the sample is quite similar to the population in terms of industry make-up. Our sample appears to be a good representation of the population, thus our findings are likely to generalize well to the population of target firms.

Model

Data on the target firm's characteristics, along with characteristics about the purchase (e.g., whether the purchase is horizontal, vertical, or conglomerate) are gathered primarily from Compustat. For each characteristic, Table 3 lists the variable name, description, Compustat mnemonic including how the variable is measured and/or defined, and the expected sign on each variable, and the hypothesis to which the variable relates. All characteristics help to explain why purchase premiums exist. Therefore, all characteristics are expected to be positively related to goodwill.

Table 3 summarizes by category the independent variables used in the analysis. Because different independent variables may capture different aspects of the categories, several hypotheses are represented by more than one variable. For example, vertical integration, horizontal integration, and geographic location all relate to customer-based characteristics (H1). Alternative specifications of the model are tested with each variable within a category separately proxying for the category, combinations of variables proxying for the category, and all variables within the category proxying for the category. Levels (dollar) variables, including goodwill, are scaled by target total assets and, similar to Palepu (1986) and Morck et al. (1990), levels variables, excluding goodwill, are averaged over the prior three years in order to minimize the effect of unusual fluctuation between years.

The following cross-sectional model is used to test the hypotheses:[3]

$$GDWL_{i*j,t1} = b_0 + b_1VI_{i,t0} + b_2HI_{i,t0} + b_3GEO_{i,t0} + b_4RD_{i,t0} + b_5IALGW_{i,t0} + b_6MT_{i,t0} + b_7TL_{i,t0} + b_8BORR_{i,t0} + b_9LEV_{i,t0} + b_{10}CASH_{i,t0} + e_{i,t1} \quad (4)$$

Where *t1* = acquisition date,

t0 = most recent annual report date prior to acquisition date *t1*,

$GDWL_{i*j,t1}$ = the book value of goodwill resulting from firm *j* purchasing firm *i*, time *t1*,

$VI_{i,t0}$ = vertical integration measure for firm *i*, time *t0*,

$HI_{i,t0}$ = horizontal integration measure for firm *i*, time *t0*,

$GEO_{i,t0}$ = geographical location measure for firm *i*, time *t0*,

$RD_{i,t0}$ = research and development expense for firm *i*, time *t0*,

$IALGW_{i,t0}$ = intangible assets less goodwill for firm *i*, time *t0*,

$MT_{i,t0}$ = managerial talent (target firm *i*, sales compared to sales in firm *i*'s industry), firm *i*, time *t0*,

$TL_{i,t0}$ = tax loss carryforward for firm *i*, time *t0*,

$BORR_{i*j,t0}$ = parent, firm *j,* borrowing rate divided by the target, firm *i,* borrowing rate, time *t0*,

$LEV_{i*j,t0}$= parent, firm *j,* debt-to-equity ratio divided by the target, firm *i,* debt-to-equity ratio, for firm *i*, time *t0*, and
$CASH_{i,t0}$= cash, firm *i*, compared to cash in firm *i*'s industry, time *t0*.

Table 3: Variable Definitions, Compustat Mnemonic, and Expected Sign by Hypothesis

Variable	Description of Item	Variable Calculation and Compustat Mnemonic	Expected Sign
H1: Customer-Based/Market-Based Characteristics			
VI	Vertical integration – primary industry code (SIC)	1 if in supplier/distributor SIC code, 0 otherwise	+
HI	Horizontal integration – primary industry code (SIC)	1 if in same SIC code, 0 otherwise	+
GEO	Geographic areas of major lines of business	1 if GEO in new area, 0 otherwise	+
H2: Technology-Based Characteristics			
RD	Research and development	(XRD+XRD[-1]+XRD[-2])/3AT	+
H3: Statutory-Based Characteristics			
IALGW	Intangibles other than goodwill	[(INTAN-GDWL)+(INTAN[-1] -GDWL [-1])+(INTAN[-2] - GDWL[-2])]/3AT	+
H4: Workforce-Based Characteristics			
MT	Managerial talent – firm sales compared to industry (market power)	SALE>=@CAVG(SALE, @SET ($C SIC=@CHGCO(COMPANY):SIC))	+
H5: Corporate Organizational/Financial Characteristics			
CASH	Cash reserves compared to industry	CHE>=@CAVG(CHE, @SET($C, SIC= ($C,SIC=@CHGCO(COMPANY):SIC))	+
TLCF	Tax loss carryforwards	[((TR/100)*TLCF) + ((TR[-1]/100*TLCF [-1])+((TR[-2]/100*TLCF[-2])]/3AT	+
BORR[c]	Borrowing rate – long term [(XINT)/((DT[-1]+DT[-2])/2) target]	[(XINT)/((DT[-1]+DT[-2])/2) parent]/	+
LEV	Debt-to-equity ratio	DTEQ parent/DTEQ target	+

All variables are related to the acquired (target) firm unless stated otherwise. Levels variable are averaged over the prior three years. Variables not found on Compustat were hand collected from published sources. SIC codes are measured at the two-digit level.

RESULTS

Descriptive Statistics

Means and ranges of values for the dependent and independent variables appear in Table 4. With the exception of goodwill, these measures are taken from the target firm prior to acquisition.

Goodwill is the amount reported by the acquirer as a result of the purchase. The number of observations for each variable ranges from 79 to 149.

Table 4: Descriptive Statistics

Levels variables (millions of dollars):	Observations	Mean	Min	Max
CASH	148	1.08	0.00	26.58
GDWL	149	1.12	0.04	8.42
IALGW	118	0.04	0.00	2.21
RD	79	0.05	0.00	0.33
TL	141	0.12	0.00	3.29
TA	149	512	1.18	10210
Ratio variables (times):				
BORR	147	5.06	0.00	545
LEV	149	29.53	-6.12	2177
SALES	148	1.54	0.00	53.22
Dichotomous variables:				
GEO	115	0.20	0.00	1.00
HI	147	0.76	0.00	1.00
VI	147	0.04	0.00	1.00

- All variables are deflated by total assets (TA); averaged over three years (except for goodwill related to purchase (GDWL) which is only deflated by total assets).

- GDWL= the book value of goodwill resulting from firm *j* purchasing firm *i*, VI = vertical integration measure for firm *i*, HI = horizontal integration measure for firm *i*, GEO = geographical location measure for firm *i*, RD = research and development expense for firm *i*, IALGW = intangible assets less goodwill for firm *i*, MT = managerial talent (target firm *i*, sales compared to sales in firm *i*'s industry), firm *i*, TL = tax loss carryforward for firm *i*, BORR = parent, firm *j,* borrowing rate divided by the target, firm *i,* borrowing rate, LEV = parent, firm *j,* debt-to-equity ratio divided by the target, firm *i,* debt-to-equity ratio, for firm *i*, and CASH = cash, firm *i*, compared to cash in firm *i*'s industry.

Multicollinearity Diagnostics

Pearson/Spearman correlation matrices are presented in Table 5. Because several correlations are statistically significant, we also calculate variance inflation factors (VIFs). A VIF in excess of 10 indicates that multicollinearity may influence least squares estimates (Neter, et al., 1996). The VIFs range from 1.17 to 2.98 implying that the results are not influenced unduly by multicollinearity in the sample data.

Table 5: Pearson (Spearman) Correlations Above (Below) the Diagonal for Goodwill Characteristics

	GDWL	VI	HI	GEO	RD	IALGW	MT	TL	BORR	LEV	CASH
GDWL	1.00	0.02 (0.81)	-0.21 (0.01)	-0.16 (0.08)	0.18 (0.10)	0.15 (0.11)	0.05 (0.55)	0.25 (0.00)	0.21 (0.01)	0.19 (0.02)	-0.01 (0.90)
VI	0.09 (0.30)	1.00	-0.29 (0.00)	0.11 (0.26)	-0.07 (0.52)	-0.04 (0.67)	0.02 (0.84)	-0.05 (0.56)	-0.02 (0.81)	-0.03 (0.73)	0.02 (0.84)
HI	-0.23 (0.00)	-0.29 (0.00)	1.00	0.00 (0.98)	0.11 (0.34)	0.05 (0.58)	-0.18 (0.02)	0.12 (0.14)	0.03 (0.70)	-0.09 (0.27)	-0.14 (0.09)
GEO	-0.15 (0.11)	0.11 (0.26)	0.00 (0.98)	1.00	-0.08 (0.54)	0.09 (0.41)	-0.12 (0.21)	0.05 (0.58)	-0.05 (0.58)	0.05 (0.60)	-0.03 (0.78)
RD	0.23 (0.04)	-0.04 (0.74)	0.02 (0.88)	0.15 (0.25)	1.00	0.41 (0.00)	-0.11 (0.33)	0.28 (0.02)	0.00 (0.98)	0.02 (0.85)	-0.09 (0.41)
IALGW	0.13 (0.17)	-0.09 (0.31)	-0.05 (0.56)	-0.12 (0.26)	0.11 (0.37)	1.00	0.05 (0.60)	0.12 (0.21)	-0.02 (0.84)	-0.02 (0.79)	-0.01 (0.93)
MT	-0.05 (0.51)	0.07 (0.43)	-0.11 (0.21)	0.02 (0.81)	-0.15 (0.18)	0.09 (0.31)	1.00	-0.06 (0.46)	-0.03 (0.75)	-0.03 (0.70)	0.46 (0.00)
TL	0.06 (0.45)	0.01 (0.94)	0.16 (0.06)	0.08 (0.39)	0.18 (0.11)	0.09 (0.32)	-0.11 (0.19)	1.00	-0.03 (0.77)	-0.04 (0.66)	-0.07 (0.39)
BORR	0.05 (0.56)	-0.10 (0.22)	0.12 (0.14)	-0.04 (0.67)	-0.11 (0.33)	-0.03 (0.74)	0.01 (0.95)	-0.13 (0.12)	1.00	-0.00 (0.99)	-0.03 (0.69)
LEV	0.06 (0.47)	-0.12 (0.15)	0.05 (0.56)	- 0.00 (0.97)	0.16 (0.16)	-0.10 (0.28)	0.09 (0.25)	-0.16 (0.06)	0.17 (0.04)	1.00	-0.03 (0.71)
CASH	-0.07 (0.40)	0.03 (0.75)	-0.21 (0.01)	0.15 (0.12)	-0.05 (0.65)	0.05 (0.56)	0.58 (0.00)	-0.04 (0.60)	-0.07 (0.43)	0.07 (0.40)	1.00

- GDWL= the book value of goodwill resulting from firm *j* purchasing firm *i*, VI = vertical integration measure for firm *i*, HI = horizontal integration measure for firm *i*, GEO = geographical location measure for firm *i*, RD = research and development expense for firm *i*, IALGW = intangible assets less goodwill for firm *i*, MT = managerial talent (target firm *i*, sales compared to sales in firm *i*'s industry), firm *i*, TL = tax loss carryforward for firm *i*, BORR = parent, firm *j*, borrowing rate divided by the target, firm *i*, borrowing rate, LEV = parent, firm *j*, debt-to-equity ratio divided by the target, firm *i*, debt-to-equity ratio, for firm *i*, and CASH = cash, firm *i*, compared to cash in firm *i*'s industry.

Regression Results for Goodwill Model

Table 6 presents the results of the model estimation. All reported regression equations also include dichotomous control variables for multiple parent purchases and the purchase year. The Breusch-Pagan test indicates that heteroskedasticity is present necessitating the use of White's heteroskedasticity-consistent estimators.[4]

Hypothesis 1 predicts that customer/market-based characteristics of the target firm are positively related to goodwill. No coefficient estimates for customer/market-based characteristics ($ß_1$, $ß_2$, $ß_3$ in Table 6) in any regression are significantly greater than zero. These results fail to provide evidence that acquirers are willing to pay a premium for the customer/market-based characteristics of target firms. Included in this category is a proxy for synergy, HI, which carries a *negative* coefficient. This result, although puzzling, is consistent with Healy, Palepu, and Ruback (1997) who found that higher synergies were related to lower premiums.

Table 6: Regression Results for All Characteristics Comprising Goodwill and Various Subsets of Characteristics Comprising Goodwill

$GDWL_{i*j,t1} = b_0+b_1VI_{i,t0}+b_2HI_{i,t0}+b_3GEO_{i,t0}+b_4RD_{i,t0}+b_5IALGW_{i,t0}+b_6MT_{i,t0}+b_7TL_{i,t0}+b_8BORR_{i,t0}+b_9LEV_{i,t0}+b_{10}CASH_{i,t0}+e_{i,t1}$

n	$\overline{R}^2$	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
All Characteristics												
51	0.37	1.15	-0.38	-0.49	-0.27	0.68	16.28	0.08	0.67	0.04	0.00	-0.03
		(2.44)	(-1.20)	(-1.04)	(-1.23)	(0.37)	(1.32)	(0.41)	(1.27)	(3.16)	(4.11)	(-0.11)
Customer-Based/Market Based Characteristics												
114	0.07	1.51	-0.09	-0.58	-0.41	-----	-----	-----	-----	-----	-----	-----
		(5.56)	(-0.18)	(-2.11)	(-2.68)							
Technology-Based, Statutory-Based, and Workforce-Based Characteristics												
63	0.05	1.56	-----	-----	-----	3.16	11.58	0.02	-----	-----	-----	-----
		(3.18)				(0.91)	(2.07)	(0.21)				
Corporate Organizational/Financial Characteristics												
138	0.17	0.95	-----	-----	-----	-----	-----	-----	1.19	0.01	0.00	0.01
		(2.92)							(1.43)	(1.49)	(9.07)	(0.45)
Technology-Based Characteristics												
79	0.00	1.45	-----	-----	-----	4.73	-----	-----	-----	-----	-----	-----
		(3.31)				(1.53)						
Statutory-Based Characteristics												
118	0.07	1.34	-----	-----	-----	-----	0.78	-----	-----	-----	-----	-----
		(3.96)					(4.74)					
Workforce-Based Characteristics												
148	0.02	1.26	-----	-----	-----	-----	-----	0.01	-----	-----	-----	-----
		(4.19)						(1.51)				

- The table shows OLS coefficient estimates and t-statistics based on White's consistent covariance estimator in parentheses.
- All level variables are reported in millions, averaged over the prior three years (except GDWL) and deflated by total assets.
- GDWL= the book value of goodwill resulting from firm *j* purchasing firm *i*, VI = vertical integration measure for firm *i*, HI = horizontal integration measure for firm *i*, GEO = geographical location measure for firm *i*, RD = research and development expense for firm *i*, IALGW = intangible assets less goodwill for firm *i*, MT = managerial talent (target firm *i*, sales compared to sales in firm *i*'s industry), firm *i*, TL = tax loss carryforward for firm *i*, BORR = parent, firm *j,* borrowing rate divided by the target, firm *i,* borrowing rate, LEV = parent, firm *j,* debt-to-equity ratio divided by the target, firm *i,* debt-to-equity ratio, for firm *i*, and CASH = cash, firm *i*, compared to cash in firm *i*'s industry.
- Values for one tailed t-tests are as follows: $t_{0.10} = 1.29$; $t_{0.05} = 1.66$; $t_{0.025} = 1.98$; $t_{0\cdot01} = 2.37$; $t_{0\cdot005} = 2.63$.

Hypothesis 2 predicts that technology-based characteristics of the target firm are positively associated with recorded goodwill. The coefficient estimated for R&D ($ß_4$ in Table 7) is not statistically significant in any specification including multiple characteristics. It becomes significant at the 0.10 level in the regression including only R&D and the control variables. Although Hirschey and Weygandt (1985) find that R&D expenditures of the current period benefit future periods, our results provide only weak evidence that acquirers are willing to pay a premium for the technology-based characteristics of target firms.

Hypothesis 3 predicts that statutory-based characteristics of the target firm are positively associated with recorded goodwill. Unlike Chauvin and Hirschey (1994) who find mixed results regarding the relation between goodwill and intangible assets, we find that the coefficient estimated for intangible assets other than goodwill ($ß_5$ in Table 6) is statistically significant and positive in all of our specifications. As such, our results provide evidence that acquirers are willing to pay a premium for the target's statutory-based characteristics.

Hypothesis 4 predicts that workforce-based characteristics of the target firm are positively associated with recorded goodwill. The coefficient estimated for workforce-based characteristics (b_6 in Table 6) is not statistically significant in any specification including multiple characteristics. It becomes significant at the 0.10 level in the regression including only managerial talent and the control variables. Thus, our results provide only weak evidence that acquirers are willing to pay a premium for the target's workforce-based characteristics.

Hypothesis 5 predicts that corporate organizational/financial characteristics of the target firm are positively associated with goodwill. The coefficient estimated for tax loss carryforward (b_7), parent borrowing rate to target borrowing rate (b_8), and parent debt-to-equity to target debt-to-equity (b_9), are significantly positive in all specifications reported in Table 6. The coefficient estimated for cash is not significantly different from zero. This result is consistent with the theory of free cash flow (Jensen, 1986) which states that excess cash is valuable only if the acquirer's management is able to put it to better use than is the target firm's management. At acquisition, the future effective use of excess cash is unproven. As a group, the coefficients estimated for corporate organizational/financial characteristics are significantly positively related to goodwill. Thus, our results provide evidence that acquirers are willing to pay a premium for the targets' corporate organizational/financial characteristics.

CONCLUSIONS AND IMPLICATIONS

This study investigates the relation between the goodwill recorded in a purchase transaction and the characteristics of the *target* firms that are acquired. The tests of the hypotheses provide insight as to which characteristics compel purchasing firms to pay a premium for the target. In summary, the data support some but not all hypothesized relations. Our results provide evidence that acquirers are willing to pay a premium for statutory-based characteristics and some corporate organizational/financial characteristics. Our results provide only weak evidence that goodwill is

positively related to technology-based characteristics and workforce-based characteristics of the targets. Finally, our results fail to show that goodwill is positively related to the customer/market-based characteristics of the target. Whether statistical significance might result from larger samples available when fewer characteristics are tested or from an omitted variables problem is unclear.

While prior research on goodwill has largely ignored the characteristics of the target firm, this study contributes to the literature by explicitly investigating the relation between the goodwill booked by the acquirer and the *target* characteristics that may explain why an acquirer paid a premium for the target. This research sets the stage for future research that further investigates goodwill utilizing a holistic approach that considers aspects of the acquirers as well as those of the target firms.

ENDNOTES

1 In accordance with SFAS No. 141, in a business combination, a customer list capable of being sold by the entity would be recognized and recorded at its estimated fair value.

2 This requires the elimination of one observation. All analyses were performed with the firm as both a target and as a parent. Because results were similar, only the analyses with the firm as a parent are reported.

3 This model also contains dichotomous control variables for multiple parent purchases and for the purchase year. There were fifteen multiple parent purchases. The number of control variables for these multiple purchases vary by model.

4 Because several models have few degrees of freedom, the White test is not feasible. The Breusch-Pagan (Neter et al., 1996) test is thus used to examine constancy of the error variances. The Breusch-Pagan test requires specification of variables related to the variance. For this model, all the variables, including the control variables and the intercept, are included in the specification. The sample statistics range from 31.43 to 100.93, far exceeding the limits ranging between 13.36 and 27.2.

REFERENCES

Amihud, Y. & B. Lev. (1981). Risk reduction as a managerial motive for conglomerate mergers. *The Bell Journal of Economics*. 12: 605 – 617.

Amir, E., M. Kirschenheiter & K. Willard. (1997). The valuation of deferred taxes. *Contemporary Accounting Research*. 14 (4): 597 – 622.

Barlev, B. (1973). Business combinations and the creation of goodwill. *Accounting and Business Research*. 304 – 308.

Berkovitch, E. & M. P. Narayanan. (1993). Motives for takeovers: an empirical investigation. *Journal of Financial and Quantitative Analysis*. 28 (3): 347 – 379.

Bradley, M., Desai, A. & E. Kim. (1988). Synergistic gains from corporate acquisitions and their division between the stockholders of target and acquiring firms. *Journal of Financial Economics*. 21: 3 – 40.

Chauvin, K. W. & M. Hirschey. (1994). Goodwill, profitability and the market value of the firm. *Journal of Accounting and Public Policy*. 13: 159 -180.

Colley, J. R. & A. G. Volkan. (1988). Accounting for goodwill. *Accounting Horizons*. March: 35-41.

Crawford, D.& T. A. Lechner. (1996). Takeover premiums and anticipated merger gains in the US market for corporate control. *Journal of Business and Accounting*. 23 (5): 807 – 829.

Doukas, J. & N. Travlos. (1988). The effect of corporate multinationalism on shareholders' wealth: Evidence from international acquisitions. *The Journal of Finance*. 43 (5): 1161-1175.

Emery, K. G. (1951). Should goodwill be written off? *The Accounting Review*. 26 (October): 560 – 567.

Falk, H. & L. A. Gordon. (1977). Imperfect markets and the nature of goodwill. *Journal of Business Finance & Accounting*. 4 (4): 443 – 462.

Financial Accounting Standards Board. (1999). Business combinations and intangible assets. *Financial Accounting Series*. Exposure draft. 201-A. September.

Financial Accounting Standards Board. (2001). Business combinations. *Financial Accounting Series*. Statement of Financial Accounting Standards No. 141. 221-C. June.

Gynther, R. S. (1969). Some conceptualizing on goodwill. *The Accounting Review*. April: 247–255.

Haw, I., V. Pastena & S. Lilien. (1987). The association between market-based merger premiums and firms' financial position prior to merger. *Journal of Accounting, Auditing, and Finance*. 2 (1): 24 – 42.

Hayn, C. (1989). Tax attributes as determinants of shareholder gains in corporate acquisitions. *Journal of Financial Economics*. 23: 121 – 153.

Healy, P. M., K. G. Palepu R. S.& Ruback. (1997). Which takeovers are profitable? Strategic or financial? *Sloan Management Review*. 38(4): 45 - 57.

Hirschey, M. & J. J. Weygandt. (1985). Amortization policy of advertising and research and development expenditures. *Journal of Accounting Research*. 23 (1): 326 – 335.

Jennings, R., J. Robinson, R. B. Thompson II & L. Duvall. (1996). The relation between accounting goodwill numbers and equity values. *Journal of Business Finance and Accounting*. 23 (4): 513 - 533.

Jensen. M. C. (1986). Presidential address: the modern industrial revolution, exit, and the failure of internal control systems. *The Journal of Finance*. 48 (3): 831 – 880.

Jensen, M. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *AEA papers & proceedings*. 76 (2):

Jensen, M. C. & R. S. Ruback. (1983). The market for corporate control. *Journal of Financial Economics*. 11: 5 – 50.

Lewellen. W. (1971). A pure financial rationale for the conglomerate merger. *Journal of Finance*. (May): 521 – 545.

Lintner J. (1971). Expectations, mergers, and equilibrium and purely competitive securities markets. *American Economic Review*. (March): 101 – 111.

Ma, R. & R. Hopkins. (1988). Goodwill—an example of puzzle-solving in accounting. *ABACUS.* 24 (1) 75 – 85.

Maquieira, C. P.; Megginson, W. L. & L. Nail. (19980. Wealth creation versus wealth distributions in pure stock-for-stock mergers. *Journal of Financial Economics*. 48: 3 – 33.

McCarthy, M. G. & D. K. Schneider. (1995). Market perception of goodwill: Some empirical evidence. *Accounting and Business Research*. 26 (1): 69 - 81.

Melnik, A. & M. Pollatschek. (1973). Debt capacity, diversification, and conglomerate mergers. *Journal of Finance*. 28 (December): 1263- 1273.

Miller, M. C. (1973). Goodwill—an aggregation issue. *The Accounting Review*. 48 (2): 280-291.

Morck, R., A. Shleifer & R. Vishny. (1990). Do managerial objectives drive bad acquisitions? *The Journal of Finance*. 45(1): 31–48.

Nelson. R. H. (1953). Momentum Theory of Goodwill. *The Accounting Review*. 28 (4): 491 – 499.

Neter, J., M. Kutner, C. Nachtsheim & W. Wasserman (1996).. *Applied Linear Statistical Models.* Chicago: Times Mirror Higher Education Group, Inc.

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

Raad, E., R. Ryan & J. Sinkey. (1999). Leverage, ownership structure, and returns to shareholders of target and bidding firms. *Quarterly Journal of Business and Economics*. 38(2): 37.

Roll, R. (1986). The hubris hypothesis of corporate takeovers. *Journal of Business*. 59 (2): 197 – 216.

Scott, J., Jr. (1977). On the theory of conglomerate mergers. *The Journal of Finance*. 32 (December): 1235 – 1250.

Smith, R. L. & J. H. Kim. (1994). The combined effects of free cash flow and financial slack on bidder and target stock returns. *The Journal of Business*. 67(2): 281 – 310.

Vincent, L. (1997). Equity valuation implications of purchase versus pooling accounting. *The Journal of Financial Statement Analysis*. 2 (4): 5 – 19.

Walker, G. T. (1938). Goodwill on financial statements. *The Accounting Review*. 13 (June): 174-182.

Walkling, R. & R. O. Edmister. (1985). Determinants of tender offer premiums. *Financial Analysts Journal*. 41 (January/February): 27 – 37.

Weston, F. & P. Halpern. (1983). Mergers and acquisitions: corporate acquisitions: A theory of special cases? A review of event studies applied to acquisitions. *The Journal of Finance*. 38 (2): 297 – 317.

Zhang, H. (1998). US evidence on bank takeover motives: a note. *Journal of Business Finance and Accounting*. 24 (7): 1025 – 1033.

THE EFFECT OF OWNERSHIP STRUCTURE ON PERFORMANCE OF HOSPITALS

Pervaiz Alam, Kent State University
Essam Elshafie, Northeastern Illinois University
David Jarjoura, Ohio State University

ABSTRACT

The purpose of this study is to investigate the effect of ownership structure on the financial performance of. We compare three groups of hospitals, classified by ownership structure: government, not-for-profit, and for-profit. Our sample is based on the data from the State of Washington hospitals over the period 1980 to 2003. Using factor analysis we summarize a large number of financial and non-financial performance measures across five dimensions: non-financial performance, profitability, capital structure, fixed assets efficiency, and liquidity. Then we apply linear mixed models analysis to investigate the effect of the ownership type on the factor scores of hospital performance. The results show that the ownership structure is associated with the level of performance, and that there are significant differences in the financial performance among the three groups of hospitals.

INTRODUCTION

The health care sector represents an important segment of the service oriented U.S. economy. As the cost of healthcare continues to rise, communities are faced with the choice of which organizational structure provides the most efficient service. Several empirical studies address the relative efficiency of hospital ownership and organizational structure. Eldenburg and Krishnan (2003) investigate the incentives and performance differences in municipal district hospitals versus private not-for-profit hospitals. They argue that the municipal district hospitals executives are paid less than those of the private not-for-profit hospitals. As a result, a selection problem arises leading to poor financial performance of district hospitals. Carter, Massa, and Power (1997) compare the operating efficiency of proprietary and non-proprietary hospitals. As a performance measure, they use the level of expenses as a function of ownership structure. They find that administrative expenses are greater for proprietary hospitals compared to non-proprietary hospitals. Many of the studies in this research area compare between two types of ownership structures of hospitals (e.g., for-profit vs. not-for-profit or private vs. public) and use limited number of financial performance measures (e.g., operating margin and expense level). In this study, we compare the performance of

governmental, for-profit, and not-for-profit hospitals using a comprehensive set of financial and non-financial measures over a period of 23 years.

The purpose of the study is to investigate the association between the ownership structure and the financial performance of hospitals. The Washington State Department of Health (WSDOH) provided us with financial and non-financial data on about 125 hospitals. These hospitals are classified into three main groups: governmental, for-profit, and not-for-profit. In turn, the governmental hospitals are classified into state, county, and district hospitals. The for-profit hospitals are classified into individual, partnership, and corporate hospitals. Finally, the not-for-profit hospitals are classified into church operated and other. The sample used in this study covers the period between 1980 and 2003. Drawing on prior research on the financial and non-financial performance of heath care organizations (e.g., Watkins, 2000 and Zeller et al., 1996), we identify and calculate annual financial and non-financial performance measures for hospitals in our sample. Then factor analysis is used to create factors that summarize the underlying characteristics of these performance measures. Five factors were identified that represent profitability, capital structure, fixed assets efficiency, liquidity, and non-financial performance. Linear mixed models analysis is used in order to properly account for the correlation structure of the data and the annual performance nested within each hospital data. Our research approach has two particular distinguishing features: 1) we use factor analysis to combine highly related variables into constructs with clearer meaning than can be obtained by using multiple measures to represent a construct and 2) we use linear mixed model to control for hospital and year effects. We believe that ignoring year and hospital effects underestimates the error variance resulting in false claims of significant differences between the hospital groups.

We found significant differences in the levels of profitability and liquidity among the three main groups of hospitals. In addition, we find significant difference across the non-financial factors. The results indicate that not-for-profit hospitals generally provide higher level of profitability and have marginally less leverage in the capital structure relative to the government hospitals. Relative to for-profit hospitals, not-for-profit hospitals rank higher in terms of profitability. The results show that the government hospitals score higher than for-profit hospitals only in terms of liquidity.

Several studies evaluate the relative efficiency of various organizational and ownership structures of hospitals. Some of these studies, however, use only a few selected financial measures. For example, Eldenburg and Krishnan (2003) use operating margin and excess income margin and the change in these two variables. Considering the complexity of the health care industry, using one or two financial measures does not provide a comprehensive evaluation of hospital performance. Other studies use a short sample period. For example, Carter et al. (1997) use data from 1989 year only. Our study analyzes several financial and non-financial measures over an extended period, 1980 to 2003. In an attempt to provide a more comprehensive analysis of the association between ownership type and hospital performance, we used factor analysis for data reduction and linear mixed model to address the correlated structure of the data.

PRIOR RESEARCH

Prior research that evaluates different types of ownership structures in hospitals focus on the objectives and relative efficiency of hospitals. Deneffe and Masson (2002) develop an empirical model to identify the objective function of not-for-profit hospitals. Their results are consistent with an objective function that places positive utility weight on both social welfare and profit. Eldenburg et al. (2004) indicate that all hospitals care about economic viability and quality of patient care, and that they, except for government hospitals, are concerned about profit margins so that they can acquire current medical technology and provide critical improvements in hospital care. Lynk (1995) tests the proposition that not-for-profit and for-profit hospitals price their hospital care differently in response to competitive structure of the market. He points out that cost of patient care, not profit margins, are most suitable for the analysis of not-for-profit hospitals' behavior. He finds that, in response to either market share or to the joint combination of market share and market concentration, private not-for-profit hospital charges are significantly lower than for-profit hospital charges.

Gulley and Santerre (1993) investigate the tax effects on the market share of the not-for-profit hospitals. They hypothesized and find that since not-for-profit hospitals are tax exempt, higher state corporate income and local property tax rate lead to a greater market share for not-for-profit hospitals. Bolton and Mehran (2006) argue that tax exempt status which the not-for-profit organizations enjoy has little economic justification. They further indicate that the effect of the market discipline resulting from competition with for-profit organizations is somewhat distorted as a result of income tax exemption enjoyed by not-for-profit organizations.

Gulley and Santerre (1993) state that while for-profit hospitals are driven by profit motive, the governmental hospitals are considered to act as "providers of last resort." The government hospital fills unmet demand for medical services. Gowrisankaran and Town (1997) develop a dynamic model to form a better methodology for predicting the effects of policy and other institutional changes upon aggregate hospital behavior and on patient welfare. In their model, there are three differences between for-profit and not-for-profit hospitals. First, the for-profit hospitals maximize their profit while the not-for-profit hospitals maximize some combination of net profit and quality of service. Second, tax codes often give not-for-profit organization tax-exempt status, while for-profit hospitals must pay property and corporate income taxes. Third, as not-for-profit hospitals cannot issue equity and their bonds are tax-exempt, the for-profit and not-for-profit hospitals two groups will have differing cost of capital. Dugan (2002) finds that not-for-profit hospitals in for-profit intensive areas are significantly more responsive to competitive changes than their counterparts in areas served by few for-profit providers. He suggests that not-for-profit hospitals mimic the behavior of private for-profit providers when they actively compete with them. Duggan (2000) examines how a hospital's type of ownership influences its response to profitable opportunities that are created by changes in government policy. He lists three different theories explaining the differences among the three types of ownerships. These theories are related to

difference in compensations schemes, types of individuals who choose to work in each types, and budget constraints.

Agency literature indicates that contracts are designed to align manager's interests with those of the principal by making the manager's compensation contingent on the output achieved. Knowing that charters of not-for-profit organizations preclude compensation arrangements that are based on firms' profit (Sloan et al., 1998), and that for-profit hospitals use incentive bonuses to increase an hospitals efficiency and effectiveness (Lambert and Larcker, 1995), one would expect that managers of for-profit hospitals would seek to maximize profits, making for-profit hospitals financially more efficient. However, prior research provides mixed evidence. Valvona and Sloan (1988) find financial efficiency, measured by return on equity and return on assets, is higher for for-profit hospitals compared to not-for-profit hospitals. Lewin et al. (1981) examine differences in cost of patients care. They attempt to identify differences in a variety of cost-related variables, between investor-owned and not-for-profit hospitals. They find that investor-owned hospitals are more expensive than not-for-profit hospitals.

Wilson and Jadlow (1982) employ a regression model to predict an efficiency divergence index. They find that proprietary hospitals are more efficient relative to other types of hospitals. Pattison and Katz (1983) use a sample of hospitals from California; they compare for-profit, not-for-profit, and government hospitals. They find that for-profit hospitals have higher administrative expenses. Lindsay (1976) finds that Veterans' administration hospitals have lower cost, and lower quality, when compared to private hospitals. Carter et al. (1997) examine the efficiency of for-profit hospitals versus not-for-profit hospitals. They find that for-profit hospitals have more administrative expenses; however, they have lower operational expenses and a lower number of employees. They also explain that the mixed results of studies like Lewin et al. (1981) and Wilson and Jadlow (1982) reflect the difficulty and complexity of determining the relative performance of hospitals by ownership structure. Sloan et al. (1998) compare cost and quality of care for Medicare patients hospitalized in for-profit hospitals contrasted with those in not-for-profit and government hospitals. They find that patients admitted to government hospitals incurred less Medicare costs, on average. The cost for for-profit and not-for-profit was similar. In terms of quality, they generally find a trivial difference among the three types of hospitals.

Kessler and McClellan (2002) find that the effects of ownership status are quantitatively important. Areas with a presence of for-profit hospitals have approximately 2.4 percent lower levels of hospital expenditures with virtually the same patient health outcomes. They conclude that for-profit hospitals have important spillover benefits for medical productivity. Comparing the efficiency of not-for-profit and government hospitals, Eldenburg and Krishnan (2003) show that the compensations level in not-for-profit hospitals are higher than those of government hospitals, creating a selection problem. The higher compensation levels of private not-for-profit hospitals, compared to government hospitals, attract talented managers resulting in a better financial performance in not-for-profit hospitals. Comparing for-profit hospitals with not-for-profit hospitals, Rosenau and Linder (2003) report that more studies have reported that not-for-profit hospitals

perform better than for-profit hospitals. He indicates that, based on a survey of 75 peer-reviewed articles on hospitals, sixty percent of the comparisons of the two groups find that not-for-profit hospital perform better than for-profit hospitals. Thirty one percent find no difference in performance between the two groups. Only eight percent of the comparisons show that for-profit hospitals perform better. Rosenau and Linder (2003) suggest that one theory as to why not-for-profit might perform better than for-profit is that these organizations are better at motivation and productivity of knowledge workers (Drucker, 1989). In addition, they do not pay taxes on income, property, and sales (Sloan, 2000), and they have access to subsidized loans (Gray, 1991b and Reinhardt, 2000).

RESEARCH QUESTIONS

Prior research provides mixed results on the relative efficiency of the three competing types of ownership structure—government, not-for-profit, and for-profit hospitals. A possible explanation is that, due to the complexity of the health care industry, using few financial performance measures may not provide a comprehensive evaluation of hospital efficiency. For example, Carter et al. (1997) use one year data on administrative expenses and number of employees to evaluate the operating efficiency in proprietary versus non-proprietary hospitals. In our study, we use a sample that covers a relatively long period of time, 1980 to 2003. In addition, we use factor analysis to summarize large number of measures into factors that underlie five characteristics of hospitals: profitability, capital structure, fixed assets efficiency, liquidity, and non-financial efficiency.

The goals and objectives of the three groups of hospitals examined in our study are different. By definition, the for-profit hospitals are motivated by profit. According to Gulley and Santerre (1993) most researchers assume that not-for-profit hospitals maximize either the quantity or the quality of the hospital services. We also assume that the not-for-profit hospitals strive to achieve these objectives without incurring losses. Finally, the government hospitals are "the providers of last resort" (Gulley and Santerre, 1993), they are opened in the underserved areas. Sine the goals and objectives of each group are unique we expect to see differences in terms of their financial and non-financial characteristics.

Knowing that managers in for-profit hospitals are compensated, at least partially, based on their financial performance while the charters of not-for-profit organizations preclude compensation arrangements that are based on profitability (Sloan et al., 1998), we expect to find for-profit hospitals to be more profitable than not-for-profit hospitals. We also expect that not-for-profit hospitals to be more profitable than government hospitals because of the selection bias (Eldenburg and Krishnan 2003).

SAMPLE

The sample consists of 125 hospitals in the State of Washington over the period from 1980 to 2003. Of these hospitals, 51 are government, 61 are not-for-profit, and 13 are for-profit hospitals. The final sample contains 2258 usable observations. An advantage of having a sample of hospitals from the same state is to control for the possible effect of different regulations. Table 1 shows descriptive statistics for the variables of interests and their description. Panel A lists non-financial measures and Panel B provides financial measures. We draw upon prior research on hospital efficiency, such as Watkins (2000), Evans et al. (2001), Eldenburg and Krishnan (2003), and Sherman (1986) to identify and define the non-financial and financial performance measures. These statistics indicate that not-for-profit hospitals are larger on average, in terms of total assets, total net revenue and number of beds licensed ($56.2 million, $2.9 million, and 208, respectively). The not-for-profit hospitals also have the highest average return on assets, the highest average return on equity and the highest average operating margin (3.99, 0.09 and 2.19%, respectively). Government hospitals are older, with average plant (hospital) age of 8.89 years, have the highest current ratio (2.2: 1), and they are the slowest in terms of collection, with average days in accounts receivable (A/R) of more than 84 days. The operating margin of the government hospitals is low (-4.30%) because the tax revenue, which is a major source of revenue to government hospitals, is deducted from net operating revenue in calculating this ratio. Excess income margin also is the lowest in government hospital (2.50%) because they provide by far the most uncompensated care (Eldenburg et al., 2004). For-profit hospitals have the highest Case Mix index [1], indicating that they treat cases that are more complex relative to the other two hospital groups. These hospitals also are faster in collecting their accounts receivable (average days in A/R = 63.8). For-profit hospitals have the lowest equity finance ratio and the highest debt to equity ratio (0.20 and 1.35, respectively), which indicates higher dependency on debt financing. They also have the highest average length of stay period (5.19 days), which may translate into higher overhead costs.

RESEARCH DESIGN AND RESULTS

Prior research on hospital performance utilizes a large number of financial and non-financial performance measures. Several studies apply factor analysis to identify fewer factors to use in hospital performance analysis (e.g., Zeller et al., 1996 and Watkins, 2000). In the current study, we also use factor analysis to identify financial and non-financial measures. We apply principal components analysis using varimax rotation and Kaiser normalization, to identify underlying factors that explain the pattern of correlation between 24 non-financial and financial measures listed in Table 1. Following Al-Eryani et al. (1990), a factor is considered when it has at least three items with loadings of 0.50 or greater. Based on this criterion, five factors were identified, namely: non-financial performance, profitability, capital structure, fixed assets efficiency, and liquidity. As shown in Table 2, these factors account for 62.15% of the total variance. Four variables loaded on

the non-financial factor: case mix, case mix adjusted admission, CMA patient days, and occupancy rate. The case mix index is a measure of the complexity of the procedures performed in certain hospitals; the case mix adjusted admission is computed as case mix index multiplied by the number of patients admitted to a hospital, and the occupancy rate that provides an efficiency measure of existing capacity utilization. CMA patient days are computed as case mix times hospital patient days. The factor loading of the CMA patient days (0.900) is the highest for this factor.

Table 1: Descriptive Statistics: Non-Financial and Financial Performance Measures (sample period 1980-2003) of hospitals in the State of Washington

	Government hospitals (46.46 % of total sample n = 1049)			Not-for-profit hospitals (46.63% of total sample n=1053)			For-profit hospitals (6.91% of total sample n = 156)		
Variable	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Panel A: Non-Financial Performance Measures									
Average age of plant	8.89	8.99	4.40	8.01	7.80	5.28	7.39	7.72	3.97
Case mix index	0.65	0.75	0.42	0.77	0.90	0.55	0.87	1.01	0.60
Case mix adjusted admission	1,881	394	4,266	5,893	3,463	6,647	2,629	1,269	2,772
CMA patient days	9,376	1,184	27,641	27,626	14,339	34,175	14,703	13,628	14,732
Length of stay	4.34	4.16	1.01	4.73	4.38	1.67	5.19	4.20	2.69
Occupancy rate	0.27	0.24	0.20	0.45	0.45	0.18	0.38	0.36	0.14
Proportion of Medicaid patients	0.16	0.12	0.11	0.15	0.13	0.10	0.18	0.11	0.17
Proportion of Medicare patients	0.50	0.50	0.16	0.45	0.47	0.15	0.43	0.49	0.21
Proportion of outpatients	0.38	0.37	0.17	0.29	0.27	0.16	0.30	0.34	0.18
Total beds licensed	81	48	92	208	179	158	127	130	51
Panel B: Financial Performance Measures									
Capital expense	0.07	0.06	0.03	0.08	.08	0.03	0.07	0.07	0.04
Cash flow to total debt	0.27	0.18	.37	0.25	0.19	0.28	0.38	0.19	0.49
Current asset turnover	3.44	3.39	0.97	3.90	3.91	.98	4.53	4.53	1.13
Current ratio	2.12	1.92	1.04	1.95	1.80	0.86	2.07	1.79	1.12
Days cash on hand	27.64	18.71	31.04	17.88	10.11	25.49	4.61	3.57	13.74
Days in A/R	84.24	80.40	24.90	69.13	67.11	28.04	63.82	57.90	51.32
Equity financing ratio	0.48	0.53	0.31	0.46	0.49	0.25	0.20	0.00	0.29
Excess income margin	2.50%	2.96%	6.71%	3.46%	3.72%	4.70%	2.54%	2.38%	7.36%
Fixed asset finance ratio	0.63	0.54	0.53	0.66	0.62	0.42	0.57	0.50	0.52
Fixed asset turnover	2.79	2.18	2.14	2.29	1.92	1.99	3.04	2.14	3.56
Long-term debt to equity	0.67	0.44	1.37	0.90	.062	1.56	1.35	0.73	2.69
Net revenue (000)	824	203	2,287	2,907	1,207	5,580	751	461	2,451
Net revenue per patient day	2,520	1,749	2,331	1,747	1,348	1,364	1,683	1,372	1,079
Operating margin	-4.30%	-2.17%	9.06%	2.19%	2.67%	4.73%	1.29%	1.78%	7.14%

Table 1: Descriptive Statistics: Non-Financial and Financial Performance Measures (sample period 1980-2003) of hospitals in the State of Washington

	Government hospitals (46.46 % of total sample n = 1049)			Not-for-profit hospitals (46.63% of total sample n=1053)			For-profit hospitals (6.91% of total sample n = 156)		
Variable	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Operating expense per patient day	2,894	1,654	5,605	2,032	1,467	2,145	2,647	1,479	3,580
Return on assets	2.57	3.32	8.27	3.99	4.34	5.88	2.78	3.01	8.93
Return on equity	0.046	0.067	0.25	0.09	0.09	0.20	(0.01)	0.04	0.29
Total assets (000)	37,543	10,001	70,293	56,276	21,747	96,185	31,475	7,421	64,046
Total asset turnover	1.08	1.03	0.40	1.03	0.96	0.36	1.34	1.30	0.51

Variable Definition
Average Age of Plant = Accumulated Depreciation / Depreciation Expense,
Capital Expense = (Interest + Depreciation) / Net Operating Expenses,
Case Mix Adjusted Admission = Case Mix index X Hospital Admission,
Cash Flow to Total Debt = (Net Income + Depreciation) / (Current Liabilities + Long-term Debt),
CMA Patient Days =Case Mix X Hospital patient days,
Current Asset Turnover = Total Operating Revenue / Current Assets,
Current Ratio = Current assets / current Liabilities,
Days Cash on hand = (Cash + Marketable Securities)/ [(Operating Expense – Depreciation)/ 365],
Days in Accounts Receivable = Net Accounts Receivable / (Patient Revenue/ 365),
Equity Financing Ratio = Fund Balance / Total Assets,
Excess income margin = [(Net Operating Revenue (Including Tax Revenue) + Non- Operating Revenue) / Total Revenue] x 100,
Fixed Assets Financing Ratio = Long-term Liabilities / Net Fixed Assets,
Fixed Asset Turnover = Total Revenue / Net Fixed Assets,
Length of Stay = Patient days/Case Mix Adjusted Admission,
Long Term Debt to Equity =Long Term Debt / Unrestricted Fund Balance,
Net Revenue Per Patient day = (Total Patient Revenue – Contractual Allowance) / Hospital Patient Days,
Occupancy Rate = Hospital Patient Days / (Number of Licensed Beds x 365),
Operating Margin = [(Net Operating Revenue – Tax Revenue) / Total Revenue] x 100,
Operating Expense per Patient Day = Total Operating Expenses / Adjusted Patient Day,
Proportion of Medicaid Patients = Medicaid Patients Days / Total Hospital Days,
Proportion of Medicare Patients = Medicare Patients Days / Total Hospital Days,
Proportion of Outpatients Days = Outpatients Days / Total Hospital Days,
Return on Assets = Net income / Total Assets,
Return on Equity = Net Income / Fund Balance, and
Total Asset Turnover = Total Operating Revenue / Total Assets.

The top and bottom 1% of the data have been eliminated to remove the effect of extreme values.

Five variables loaded on the profitability factor (cash flow to total debt, excess income, return of assets, return on equity, operating margin). These variables are traditional profitability measures that have been used in prior health care research. The return on assets (ROA) has the highest factor loading for the profitability factor (0.931). The three variables that loaded on the capital structure factor are measures of the degree of the dependency on debt financing relative to equity financing (equity financing, fixed assets finance ratio, long term debt to equity).

The three variables that loaded on the fixed assets efficiency factor measure the amount of revenue generated relative to the assets employed and the relative weight of the assets-related-expenses to the total operating expenses (capital expense, fixed assets turnover, total assets turnover). Finally, the variables that loaded on the liquidity factor measure the hospital's ability to meet its short term obligations and its handling of the available liquidity (current asset turnover, current ratio, days cash on hand). These five factors are then used as dependent variables in equation (1) and as control variables in equations (2) below (non-financial factor, profitability factor, capital structure factor, fixed assets efficiency factor, liquidity factor).

To compare performance across hospital ownership types, we used a linear mixed model to take account of the fact that multiple years of hospitals may be correlated. Since hospitals are classified into the three groups, ignoring hospital effects (treating the multiple year data set as if all years are independent) would give too small of an error variance for group comparisons[2]. In applying this method we consider the fixed effects of the type of the hospital and the random effects of both the individual hospital and the year in addition to the effect of the interaction between the hospital effect and the year effect. The equation (1) below was estimated for each of the five factors separately. The independent variables are various measures which affect performance of a hospital.

Equation (1):

$$Factor_{itg} = group_g + \beta_1\, Case\ mix\ index_{it} + \beta_2\, Number\ of\ licensed\ beds_{it} + \beta_3\, Proportion\ of\ Medicare\ patients_{it} + \beta_4\, Proportion\ of\ Medicaid\ patients_{it} + \beta_5\, Proportion\ of\ outpatients_{it} + hospital_i + year_t + year{*}group_{ig} + \epsilon_{itg}$$

where i refers to hospital i (nested in group g), subscript t refers to year, group is the effect of being the g^{th} ownership type group (g=1,2,3), β_1 to β_5 represents the coefficients for the five covariates. They are included in the model to correct for confounding bias and to add precision for comparison among groups. The Case Mix index is a measure of the severity of cases treated, included to control for the complexity of services provided. The more complex the case gets, the more cost is expected to be incurred. Phillips (2003) suggests that controlling for this factor places hospitals on a level playing field so that valid financial comparisons can be made[3]. The number of licensed beds is a measure of the available bed capacity in hospitals. Proportion of Medicare, Medicaid, and outpatients to total patients, are included to control for relative profitability of different types of patients. After 1983, Medicare paid hospitals for inpatient services at fixed rates (Eldenburg and Kallapur, 1997). Nutt and Cleverley (1984) examine a number of performance measures and find that the percentage of Medicaid revenue is significantly correlated with hospital bond ratings. Eldenburg and Krishnan (2003) indicate that margins are lower in hospitals with higher proportions of outpatients. They explain that outpatients usually generate less revenue than inpatients and are less likely to be insured because Medicare requires hospitals to treat all emergency room patients, regardless of their ability to pay. The three random effects (hospital, year, and year * group) are used to account for the correlated error structure of the data.

Table 2: Factor loadings of non-financial and financial factors					
Financial and Non-Financial Factors and their Determinants	Factor Loading	Communality	Eigenvalue	Percentage of Variance	Cumulative Variance
I. Non-financial Factor:			4.649	18.60	18.60
1- Case Mix	0.759	0.670			
2- Case Mix Adjusted	0.883	0.879			
3- CMA Patient Days	0.900	0.859			
4- Occupancy Rate	0.861	0.892			
II. Profitability Factor:			4.342	17.37	35.96
1- Cash Flow to Total Debt	0.610	0.734			
2- Excess Income	0.928	0.905			
3- Return of Assets	0.931	0.903			
4- Return on Equity:	0.826	0.714			
5- Operating Margin	0.731	0.691			
III. Capital Structure Factor:			2.689	10.76	46.72
1- Equity Financing	0.870	0.898			
2-- Fixed Assets Finance Ratio	-0.876	0.794			
3- Long Term Debt to Equity	-0.684	0.539			
IV. Fixed Assets Efficiency Factor:			1.989	7.96	54.68
1- Capital Expense	-0.613	0.686			
2- Fixed Assets Turnover	0.899	0.864			
3- Total Assets Turnover	0.896	0.840			
V. Liquidity Factor:			1.866	7.47	62.15
1- Current Asset Turnover	-0.879	0.889			
2- Current Ratio	0.694	0.939			
3- Days Cash on Hand	0.856	0.803			
The extraction method is principal components analysis and the rotation method is Varimax with Kaiser Normalization. See Table 1 for definition of variables.					

In equation (2) below, we examine differences in profitability across the three groups of hospitals using independent variables from equation (1) and non-financial factor, capital structure factor, fixed assets efficiency factor, and liquidity factor.

Equation (2): $Profitability\ Factor_{itg} = group_g + \beta_1\ Non\text{-}financial\ factor_{it} + \beta_2\ Capital\ structure\ factor_{it} + \beta 3\ Fixed\ assets\ efficiency\ factor_{it} + \beta_4\ Liquidity\ factor_{it} + \beta_5\ Number\ of\ licensed\ beds_{it} + \beta_6\ Proportion\ of\ Medicare\ patients_{it} + \beta_7\ Proportion\ of\ Medicaid\ patients_{it} + \beta_8 Proportion\ of\ outpatients_{it} + hospital_i + year_t + year*group_{ig} + \epsilon_{itg}$

Table 3 shows the results of testing for significant differences between pairs of the three hospital groups using equation (1). The estimate column provides the estimated difference between the hospitals. For the non-financial performance factor, the non-for-profit group was significantly higher than the government group with p<0.0001. The results also show that the government hospitals score marginally higher than the not-for-profit group in terms of capital structure (p = 0.06), which means that government hospitals are less dependent on debt financing relative to the not-for-profit-hospitals. On the other hand, no significant differences were found in the fixed assets efficiency factor. For the liquidity factor, the coefficient estimate for government hospitals is higher than for-profit hospitals (p=0.03). For the profitability factor, the non-for-profit group hospital is significantly higher than both for-profit hospitals (p < 0.01) and government hospitals (p=0.04). These results indicate that the not-for-profit hospitals objective is not to just to avoid incurring losses, but to provide earnings[4]. The for-profit hospitals do not show superiority over the other two groups along any of the five factors.

Table 3: The Results of Linear Mixed Modeling Pairwise Comparison among the Three Groups of Hospitals, Profit, Not-for-Profit, and Government, in Terms of Five Factors Underlying their Non-Financial Performance, Profitability, Capital Structure, Fixed Assets Efficiency, and Liquidity.

$Factor_{itg} = group_g + \beta_1\ Case\ mix\ index_{it} + \beta_2\ Number\ of\ licensed\ beds_{it} + \beta_3\ Proportion\ of\ Medicare\ patients_{it} + \beta_4\ Proportion\ of\ Medicaid\ patients_{it} + \beta_5 Proportion\ of\ outpatients_{it} + hospital_i + year_t + year*group_{ig} + \epsilon_{itg}$ (1)

Dependent variable	Label	Coefficient Esitmates	t-value	p-value
Profitability				
	Government vs. Not-for-Profit	-0.45	-3.45	*<0.01*
	Government vs. For-Profit	0.23	0.74	0.47
	Not-For-Profit vs. For Profit	0.68	2.16	*0.04*
Capital structure				
	Government vs. Not-for-Profit	0.36	1.94	*0.06*
	Government vs. For-Profit	0.21	0.48	0.64
	Not-For-Profit vs. For Profit	-0.15	-0.34	0.74
Fixed assets efficiency				
	Government vs. Not-for-Profit	0.07	0.31	0.76
	Government vs. For-Profit	-0.42	-0.75	0.46
	Not-For-Profit vs. For Profit	-0.49	-0.88	0.39
Liquidity				
	Government vs. Not-for-Profit	0.18	1.21	0.24
	Government vs. For-Profit	0.80	2.22	0.03
	Not-For-Profit vs. For Profit	0.62	1.71	*0.10*
Non-financial				
	Government vs. Not-for-Profit	-1.01	-5.20	*<0.0001*
	Government vs. For-Profit	-0.47	-0.99	0.33
	Not-For-Profit vs. For Profit	0.54	1.14	0.26

The top and bottom 1% of the data have been eliminated to remove the effect of extreme values.

In Table 4 we reexamine the differences in profitability among the three hospital groups using equation (2). The results show that even after controlling for non-financial, capital structure, fixed assets efficiency, and liquidity factors, the not-for-profit hospitals still rank higher relative to the other two hospital groups in terms of profitability.

Table 4

The Results of Linear Mixed Modeling Pairwise Comparison among the Three Groups of Hospitals, Profit, Not-for-Profit, and Government, in Terms of Profitability after Controlling for other Factors Underlying the Financial and Non-Financial Characteristics.

Profitability Factor$_{it}$ = group$_g$ + β_1 Non-financial factor$_{it}$ + β_2 Capital structure factor$_{it}$ + β_3 Fixed assets efficiency factor$_{it}$ + β_4 Liquidity factor$_{it}$ + β_5 Number of licensed beds$_{it}$ + β_6 Proportion of Medicare patients$_{it}$ + β_7 Proportion of Medicaid patients$_{it}$ + β_8 Proportion of outpatients$_{it}$ + hospital$_i$ + year$_t$ + year*group$_{ig}$ + ϵ_{itg} (2)

		Coefficient		
Performance Factor	Label	Esitmates	t-value	p-value
Model (2) Profitability Factor	Government vs. Not-for-Profit	-0.50	-3.14	*<0.01*
	Government vs. For-Profit	0.30	0.81	0.42
	Not-For-Profit vs. For Profit	0.80	2.16	*0.04*

We did not attempt to control for Type I error across these comparisons in the declarations of significance above, because of the exploratory nature of this study. However one could apply a standard p-value adjustment procedure to correct for the multiple comparisons (3 hospital groups) and the multiple variables (5 explanatory variables). When we use Bonferroni inequality to control for Type I error, the critical p-value needed is 0.0033=0.05/15. This supports the claim that there is less than a 1% chance of falsely declaring one or more comparisons significant among the 15 comparisons. With this strict criteria, two comparisons can be confidently claimed to be significant, namely not-for-profit hospitals are better than government hospitals on non-financial and profitability performance dimension. In addition, the larger p-value for the not-for-profit versus for-profit hospitals comparison in Table 4 on profitability (p=0.034) can be attributed to the smaller number of for-profit hospitals in the sample lead to inadequate power of the regression. Also, note that the difference between the not-for-profit and for-profit hospitals is larger than not-for-profit and government hospital differences (-0.50) on profitability (see Table 4).

CONCLUSIONS

The purpose of this study is to examine the difference in the levels of financial and non-financial performance among three groups of hospitals classified based on their ownership structure. Based on a sample of 125 hospitals from the State of Washington over the period of 1980 to 2003,

we investigate whether there are significant differences in performance among three groups of hospitals: government, for-profit, not-for-profit. Using factor analysis, we were able to summarize large number of financial and non-financial measures into five factors underlying hospital performance. These factors are labeled as profitability, capital structure, fixed assets efficiency, liquidity, and non-financial. We apply linear mixed models to account for the both the fixed effect of the group membership and the random effect of the individual hospital and the year. The results indicate that not-for profit hospitals are more profitable relative to for-profit and government hospitals even after controlling for other factors reflecting other financial and non-financial measures. Future studies need to use different samples from other states to evaluate the generalizability of these results.

ENDNOTES

[1] The case mix index, assigned to each provider by Medicare, measures the complexity and severity of cases treated (Phillips 2003).

[2] We used Proc Mixed procedure for performing linear mixed analysis (see SAS Proc Mixed v. 9.1, SAS Inc. Cary NC, 2004).

[3] When the non-financial factor is used as the dependent variable the case mix index was not included as a covariate.

[4] Upon discussing these results with some of the health care industry professionals, they indicated that a possible explanation for the not-for-profit hospitals higher profitability is that the charitable objective of some of the not-for-profit organizations that run hospitals is not to provide affordable health care services; rather they aim to generate earnings from the hospital to spend on other charitable activities.

REFERENCES

Al-Eryani, M. F., P. Alam & S. Akhter. (1990). Transfer pricing determinants of U.S. multinationals. *Journal of International Business Studies* 21 (3): 409-425.

Bolton, P. & H. Mehran. (2006). An introduction to the governance and taxation of not for-profit organizations. *Journal of Accounting and Economics,* 41(3): 293-305.

Carter, R. B., L. J. Massa & M. L. Power. (1997). An examination of the efficiency of proprietary hospital versus non-proprietary hospital ownership structures. *Journal of Accounting and Public Policy 16* (1): 63-87.

Deneffe, D. & R. Masson. (2002). What do not-for-profit hospitals maximize? *International Journal of Industrial Organization* 20 (4): 461–492.

Drucker, P. F. (1989). What business can learn from nonprofits. *Harvard Business Review* 67 (4): 88–93.

Dugan, M. (2000). Hospital ownership and Medical public spending. *The Quarterly Journal of Economics* 11 (4): 1343-1373.

Duggan, M. (2002). Hospital market structure and the behavior of not-for-profit hospitals. *RAND Journal of Economics* 33 (3): 433-446.

Eldenburg, L. & R. Krishnan. (2003). Public versus private governance: A study of incentives and operational performance. *Journal of Accounting and Economic* 25 (August): 377-404.

Eldenburg, L., B. E. Hermalin, M. S. Weisbach & M. Wosinska. (2004). Governance, performance objective and organizational form: Evidence from hospitals. *Journal of Corporate Finance* 10 (4): 527- 548.

Evans, J. H., Y. Hwang & N. J. Nagarajan. (2001). Management control and hospital cost reduction: Additional evidence. *Journal of Accounting and Public Policy* 20 (1): 73-88.

Gardiner, L.R., S.L. Oswald & J.S. Jahera Jr. (1996). Prediction of hospital failure. *Hospital and Health Services Administration* 41 (4): 441-460.

Gowrisankaran, G. & R. J. Town. (1997). Dynamic equilibrium in the hospital industry. *Journal of Economics and Management Strategy* 6 (1): 45-74.

Gray, B. H. (1991b). Why nonprofits? Hospitals and the future of American health care. *Frontiers of Health Services Management* 8 (4): 3–32.

Gulley, O. D. & R. E. Santerre. (1993). The effect of tax exemption on the market share of nonprofit hospital. *National Tax Journal* 46 (4): 477-486.

Kessler, D. P. & M. B. McClellan. (2002). The effects of hospital ownership on medical productivity. *Rand Journal of Economics* 33 (Autumn): 488-507.

Lambert, R. A. & D. F. Larker. (1995). The prospective payment system, hospital efficiency, and compensation contract for senior-level hospital administrators. *Journal of Accounting and Public Policy* 14 (1): 1-31.

Lewin, L. S., R. A. Derzon & R. Margulies. (1981). Investor-owned and non-profits differ in economic performance. *Hospitals* 55 (July): 52-58.

Lindsay, C. (1976). A theory of government enterprise. *Journal of Political Economy* 48 (5): 1061– 1077.

Lynk, W. (1995). Nonprofit hospital mergers and the exercise of market power. *Journal of Law and Economics* 38 (2): 437-462.

Mick, S.S. & C.G. Wise. (1996). Downsizing and financial performance in rural hospitals. *Health Care Management Review* 21 (2): 16–25.

Nutt, P.C. & W.O. Cleverley. (1984). Hospital bond rating and its managerial implications. *Health Services Research* 19 (5): 615-637.

Pattison, R. V. & H. M. Katz. (1983). Investor-owned and not-for-profit hospitals: A comparison based on California data. *The New England Journal of Medicare* 309 (6): 347- 353.

Phillips, J. F. (2003). The dilemma of valuing not-for-profit hospitals- Is free cash flow the answer? *Journal of Accounting and Public Policy* 22 (4): 347-374.

Reinhardt, U. E. (2000). The economics of for-profit and not-for-profit hospitals. *Health Affairs* 19 (6): 178–186.

Robinson, J.C. & C. S. Phibbs. (1989). An evaluation of Medicaid selective contracting in California. *Journal of Health Economics* 8 (4): 437–455.

Rosenau, P. V. & S. H. Linder. (2003). Performance evaluation of for-profit and nonprofit U.S. hospitals since 1980. *Nonprofit Management and Leadership* 13 (4): 401-423.

Sherman, H. D. (1986). Financial reporting: Interpreting hospital performance with financial statement analysis. *The Accounting Review* 61 (3): 526-550.

Sloan, F. A. (2000). Not-for-profit ownership and hospital behavior. In A. J. Culyer and J. P. Newhouse (eds.), *Handbook of Health Economics*. Amsterdam: Elsevier Science.

Sloan, F.A., G. A. Picone, D. H. Taylor Jr. & S. Y Chou. (1998). Hospital ownership and cost and quality of care: Is there a dime's worth of difference? Working paper, National Bureau of Economic Research.

Valvona, J. & F. A. Sloan. (1988). Hospital profitability and capital structure: A comparative analysis. *Health Services Research* 12 (3): 343-357.

Watkins, A. N. (2000). Hospital financial ratio classification patterns revisited: Upon considering nonfinancial information. *Journal of Accounting and Public Policy* 19 (1): 73-95.

Wilson, G. W. & J. M. Jadlow. (1982). Competition, profit incentives, and technical efficiency in the provision of nuclear medicine services. *Bell Journal of Economics* 13 (2): 472-482.

Zeller, T.L., B. B. Stanko & W. O. Cleverley. (1996). A revised classification pattern of hospital financial ratios. *Journal of Accounting and Public Policy* 15 (2): 161-182.

AN EMPIRICAL TEST OF AN IPO PERFORMANCE PREDICTION MODEL: ARE THERE "BLUE CHIPS" AMONG IPOS?

John Miller, Sam Houston State University
Robert Stretcher, Sam Houston State University

ABSTRACT

An earlier study of 563 firms which issued IPOs during 1997 identified and estimated a three-stage algorithm in which basic accounting variables and indices available at the time of the IPO were found to predict mean annual wealth appreciation from buy-and-hold stock ownership for the ensuing three years. Firm size predicted membership in the middle sixth and seventh deciles; sales, receivables turnover, and retained earnings per assets predicted the top quintile; current debt and selling costs predicted the lowest quintile. Since February 2001 market trends have been generally negative. The current paper confirms the earlier model despite negative currents.

PURPOSE OF THIS STUDY

An earlier investigation (Miller, 2003) uncovered a non-linear and, indeed, non-metric anomaly in the joint distributions of the wealth appreciation of companies with new initial public offerings and certain accounting data made public at or around the date of the offering. The earlier study was purely exploratory and consisted of specifying the model and estimating the parameters of a three-stage prediction scheme. The model was able to predict approximately three-fourths of the firms correctly into three segments of wealth appreciation. The three segments were the "MID" comprised of the sixth and seventh deciles, "TOP" or the top quintile, and "LOW" or the bottom quintile. It is the objective of this study to evaluate the performance of the model in the face of the generally poor market conditions of the two years immediately posterior to model construction (March, 2001 to July 2003).

INTRODUCTION

It is not rare to find examples of data mining in the literature relating financial data to stock market and general business performance. Even the most influential of the early papers on company failure prediction (e.g., Beaver, 1967, Altman, 1968, and Edminster, 1972) might be accused of too-enthusiastic opportunism by their use of repeated analyses (one suspects) until a statistically significant formulation appeared. And, to make matters worse, sample sizes were very small and

drawn as convenience samples rather than probability samples. As is apparent from these cautionary examples, data mining is not always a complementary term. It is also called "data dredging" or the over-working of data, and is a natural result of the statistician's desire to do a thorough job. It may be said that the goal of any statistical analysis is to uncover statistical significances. See Fisher (1986) for a broader discussion of the tensions between the statistician and his client. There is also a careful discussion of the problem in the paper and subsequent comments in Chatfield (1995). Chatfield underscores the potentials for disaster whenever a model is uncovered and fit to a set of data, and then tested on the same set. This is especially true in the cases of step-wise regression and time series analysis. While this is not a novel idea, he goes further to argue that split sample designs are also suspect and that models should preferably be tested on data gathered at another time. Only then can his "model selection biases" be removed. More generally, it can be argued that there are two stages in any kind of scientific enterprise. Tukey (1977) has developed a broad range of powerful "exploratory data" tools to assist the researcher in uncovering explanatory models. But he would agree that there is still a need for "confirmatory" analysis (Tukey, 1980). Good scientific procedure calls for such confirmation not to come from the model source, but from independent investigators operating in other sites on related, but not identical, datasets. The approach of this paper is strictly "exploratory" and the confirmatory phase will be left as a follow-up exercise.

As part of a fundamental reflection on the theoretical underpinnings of the statistical analysis, Hand (1996) has expanded on the opening provided by Velleman and Wilkinson (1993), who were criticizing the psychophysicist Stevens' (1951) data measurement scale hierarchy (nominal, ordinal, interval and ratio) that has become almost routinely accepted in much of scientific work, especially the social sciences and business research. Hand argued that the traditional approach to science used a "representational measurement theory" in which the data are integral parts of mathematical models of empirical theories and are direct attempts to "formulate properties that are observed to be true about certain qualitative attributes" (foreword to the monumental Foundations of Measurement trilogy, Krantz et al., 1971, Suppes et al,. 1989, Luce et al., 1990 quoted in Hand, 1996) This is the dominant assumption used by most scientists in their work and is at least a century old. Later, as physicists became troubled by such difficulties as those caused by the dual nature of light, physicists began to relax the relationship between their data and the real world. The development of "operational measurement theory" is traced by Hand to Bridgman (1927) and is a shift in the focus of the measurement theory from the empirical to the mathematical construct being used to model that reality. In this case, the emphasis is on how that model determines the properties of the data measurement scale. It is exemplified in the elaborate models of latent variables and structural equations used in the social sciences. There the models are less a picture of some external reality and more of a prediction scheme. Now the role of the statistician is merely to insure that the assumptions about the data structure do not violate that model, not some underlying reality. The responsibility for connection between the model and external reality is entirely that of the social scientist, not the data analyst.

There was probably a time when accounting data was (reasonably) thought to be representational. The representational approach is still exemplified in the work of the Banque de France (Bardos, 2000) in which classical Fisher linear discriminant analysis is used to forecast company disasters. But, it is becoming more and more apparent that such reliance upon the external reality of bookkeeping data is not warranted. This relaxed approach is that of Zopoundis (1999), for example. For the purposes of this analysis we will not assume that the SEC-reported numbers are fundamentally precise reflection of a company's situation, but we will assume that the data can be relied on for direction and relative size. That is, for most of the subsequent analysis we assume qualitative rather than quantitative scaling.

The model used to begin this analysis was that of correspondence analysis. It is one of many statistical procedures which have as their *raison d'etre* the analytic development of a quantitative re-scaling from data which are assumed only to be nominal or ordinal to begin with. One popular model is that of the "tolerance" distribution of Cox and Oakes (1984) and McCullagh (1980). Correspondence analysis has a long history rooted in the work of Fisher (1940) in the middle of the last century; however, it is certainly not the only possible analytic procedure. There are many possible models that are used to rescore the rows and/or columns of a contingency table. These models can have either no conditions on the scores for the rows and/or columns (unrestricted models), or it is possible to require that either the rows or the columns or both must be ordinal in nature (restricted models). On the one hand, Goodman (1981, 1986) has developed his R, C and RC models. The latter is shown (Ritov and Gilula, 1991) to be equivalent to:

$$P_{ij} = \alpha_i \beta_j \exp(\gamma \mu_i \nu_j) \qquad (1)$$

where Gamma is the coefficient of intrinsic association; the sets of parameters Alpha and Beta are the scores to be "optimized" by maximizing; Mu and Nu are nuisance parameters. The rescorings are centered to zero and scaled to one so that, according to Goodman, they can be compared to the results of correspondence analysis, in which the same standardizing takes place. Gilula, Kreiger, and Ritov (1988) show that this is a model in which entropy in an information theory sense is being maximized

Procedures for estimation of the parameters in the RC model have been developed by Goodman in the unrestricted case. The R and C association models in the restricted case were solved by Agresti, Chuang and Kezouh (1987) and the RC model by Ritov and Gilula (1993).

The correspondence model used in this paper can be expressed as:

$$P_{ij} = P_{i.} P_{.j} (1 + \lambda \varepsilon_i \delta_j) \qquad (2)$$

where Lambda is the "coefficient of stochastic extremity" (Gilula, Kreiger, and Ritov, 1988), the sets of parameters Epsilon and Delta and are the scores to be "optimized" by maximizing, and the

P's are marginal proportions. "Stochastic extremity" is reference to the cumulative distributions of the rows (columns) which are maximally distanced by this procedure. The coefficient Lambda is a monotonic correlation in the sense of Kimeldorf and Sampson (1978), which they define as the supremum of correlation coefficient over all possible monotonic functions of the two variables. Perhaps the most interesting aspect of monotonic correlation is in its relation to statistical independence. Unlike the case for an ordinary Pearson correlation coefficient, when a monotonic correlation is zero, then the variables are independent. The optimization solution in the unrestricted case of this can be traced at least as far back as Fisher (1940) and even Hotelling (1933, 1936), and can be easily derived through the singular value decomposition of a certain matrix (Hirshfield, later changed to Hartley, 1935). The latter parameter, Lambda, also has the felicitous meaning of a canonical correlation. A further appealing property of the correspondence model (and the RC model, too, for that matter) is that the data are "stochastically ordered" in the sense that if the scores are ordered, then the conditional cumulative probabilities over those scores are similarly ordered (Ritov and Gilula, 1993).

The differences between these two approaches, correspondence analysis and Goodman's RC model, to constructing the restricted ordinal scales will in general result in similar scale values and similarly interpreted measures of association – so long as the association between the pair of variables is "weak." (This is an observation by Goodman, 1981, for unrestricted models extended to restricted models by Ritov and Gilula, 1993.) This is the situation in most social sciences and business applications, and is certainly true for the properties under investigation here. In fact, under the commonly-held "market efficiency" presumption there should be no correlation at all.

For this analysis, the particular algorithm is not that from Ritov and Gilula (1993) in which they reparameterize the scales via a latent variable approach and then use the EM algorithm to optimize the scores. This analysis follows the venerable Benezecri (1973) to Gifi (1990) track which utilizes an "alternating least squares" optimization due to Young, de Leeuw and Takane in the 1970s (see de Leeuw, 1993). The specific implementation of this is found in the routine "Optimal Scoring" developed for the Statistical Package for the Social Sciences (SPSS) by Meulman (1992, but also the technical annotation for the SPSS routine).

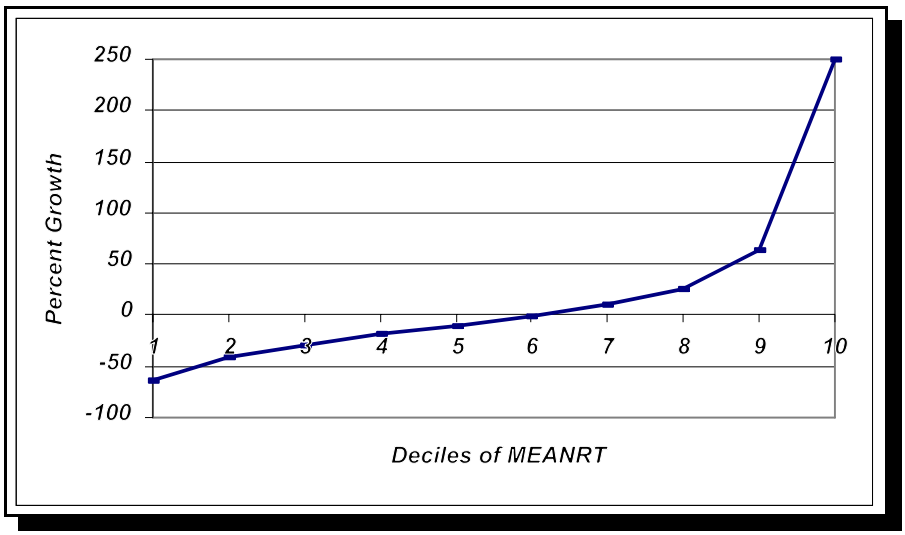
However, we actually began this analysis not at the ordinal level, but without any assumptions beyond categorization. Each of the firms' predictor variables was reduced to deciles and submitted to a correspondence analysis. (SPSS "Optimal Scoring") The use of deciles is common in finance literature (e.g., Lakonishok, Shleifer and Vishny, 1994) and it is the basic beginning data structure of this paper's analysis. Later, investigation was made into other "n-tiles" (from quintiles up to "20-tiles") only to find no real difference between the correspondence analysis results. To get a picture of the scale, and unusual nature of this financial data, here are the means of the deciles for the variable measuring the average over several years of the 12-month wealth appreciation (MEANRT). This will be the primary response variable for the subsequent analysis, and a major goal will be to rescale this to a manageable metric. While a nearly linear pattern can be seen over the middle seven deciles, the first and last two deciles break that pattern. The highest

decile had an average return four times that of the ninth decile. Note also that half of the deciles had average performances which returned either no gain or, more likely, a loss to those who held the stocks for a year.

This variable is defined by Compustat as: "The Total Return concepts are annualized rates of return reflecting price appreciation plus reinvestment of monthly dividends and the compounding effect of dividends paid on reinvested dividends "(Research Insight, 2001).

The first set of regressions were bivariate analyses of the deciles for each of the 25 financial scales and ratios versus the mean annual wealth aggregate for the period January 1998 (for those companies that went public very early in 1997) to February, 2001. Each of the 25 predictor variables was selected because it had been considered in earlier research and was available at or near the time of issuance of the IPO. (The results of the nominal-scaled analysis showed that the predictor variables could be reasonably approximated as ordinal without much loss of correlation, so the 25 optimal scaling analyses were re-run forcing an ordinal restriction on the predictors but not on the response). The table below and the subsequent graph show an example of the results of those nominal-ordinal analyses. In general, the relationships between the predictors and response were not strong, but for many they were not insignificant.

Figure 1: Mean Annual Wealth Appreciation

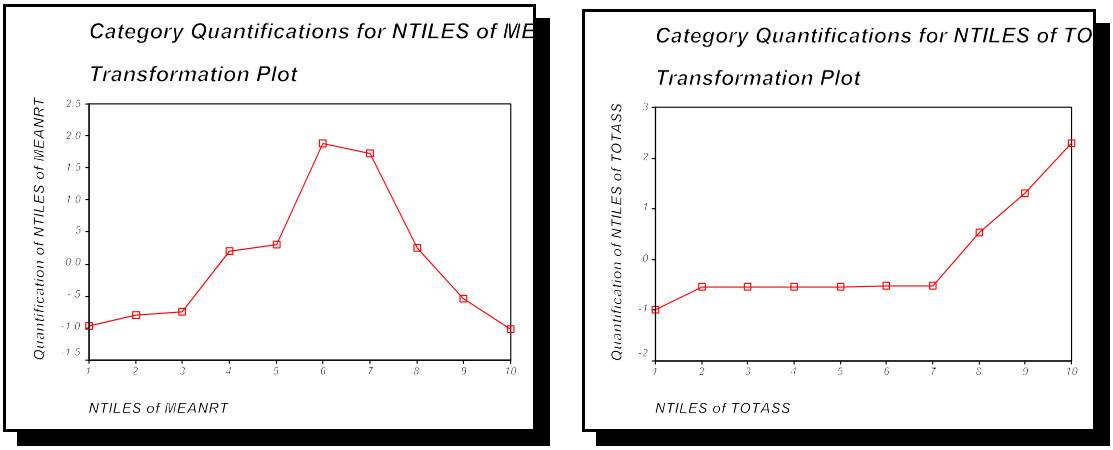


One is struck immediately by the respectable performance of such variables as TOTASS (total assets, r = .355), TOTLIAB (total liabilities, r = .344), EBITDA (earnings before taxes, r = -.343), OPER.AT (operating income to assets ratio, r = -.328), NETPROFM (net profits, r = -.319), RETE.AT (retained earnings to assets ratio, r = -.310), and SALES (net sales, r = -.306). These correlations are, of course, only potentials. They are the maxima found by a process designed to adjust the response and predictor measures (row and column scores) monotonically until such

maxima are achieved. (Recall that ordinary Pearson correlation coefficients are also maxima derived from a process of optimization over all possible linear relations.) But, these are high enough to be encouraging. Note also the valid sample sizes. No attempt was made to eliminate any special classes of businesses (REITs, financial institutions, etc.); if they reported data, they were included. And, for many of these variables, 90% or more of the 563 total firms did report the predictor variable. (See Table 1 below.)

Below are the graphs relating the original "raw" ordinal integer scores to the rescoring values ("quantifications") for an example variable which gave rise to the correlations above. There is a persistent non-linear pattern in the MEANRT (average wealth appreciation) response scorings. It strongly suggests that one end of the predictor variable scale (the right end is indicated by the sign on the correlation coefficient in the table above) is related to middling performance at or above the median ("above average firms") and that the variable is less able to discriminate between those at the extremes. The top decile and the bottom three response deciles receive nearly identical scores.

Figure 2: Optimal Rescorings



This example is only one of the very clear pictures resulting from the pairings of wealth appreciation ("MEANRT") with each of the 25 predictors. The rescoring for total assets (TOTASS - chart on right) puts the bottom 70% of the firms' assets at virtually the same score, then distinguishes between the top three deciles. The response curve (MEANRT - chart on left) has the interior sixth and seventh deciles well above the others. It would appear that the larger three-tenths of companies (in terms of assets size) tend to be in the sixth and seventh decile (in terms of wealth appreciation). This pair of deciles will be called the "MID" group.

Almost half (44%) of the middle wealth group were among the largest companies (when they went public), while only about 13% of the remaining companies were in that largest size group. It is apparent that the rescoring found by the procedure and charted above, which has the effect of equating the lowest seven deciles, leads to a linear pattern relating the two variables:

Table 1: Optimal Scoring Correlations			
Predictor	Valid Data	Correlation	R^2
EBITDA	515	-0.343	0.118
SALES	532	-0.306	0.094
COSTSALE	532	-0.254	0.064
SELLCOST	442	0.290	0.082
INTEXP	504	-0.299	0.090
CURASTOT	469	0.238	0.045
INTANG	556	-0.179	0.032
TOTASS	540	0.355	0.125
DEBTCURR	507	-0.225	0.051
TOTCLIAB	476	0.241	0.058
TOTLTDBT	540	0.276	0.076
TOTLIAB	540	0.344	0.118
DEBTEBIT	515	-0.262	0.068
LIABNETW	540	-0.201	0.041
EBITASS	515	-0.279	0.078
CURRENTR	468	0.208	0.043
QUICKRAT	474	0.206	0.043
RECTURN	489	-0.209	0.044
TOTASST	507	-0.176	0.031
CASHTURN	505	-0.269	0.073
NETPROFM	525	-0.319	0.101
WCAP.AT	467	0.186	0.034
RETE.AT	534	-0.310	0.096
OPER.AT	439	-0.328	0.108
SALES.AT	509	-0.181	0.033
Source: Compustat estimates of mean wealth appreciation for 563 IPOs in 1997. Input data were decile memberships and were not forced to be ordinal.			

Stage 1

Having discovered the (relative) isolation of the MID Group, recourse was made to a logistic regression to develop a prediction scheme for this middle group. As compared to discriminant function analysis, Press and Wilson (1978) showed that logistic regression is to be preferred,

primarily because the latter is better equipped to handle non-normal predictors. All of the variables identified above as having high potential correlations were tried out. The best prediction equation (after a variation of the all possible subsets regression paradigm) turned out to be:

Figure 3: MID Group Membership Rates

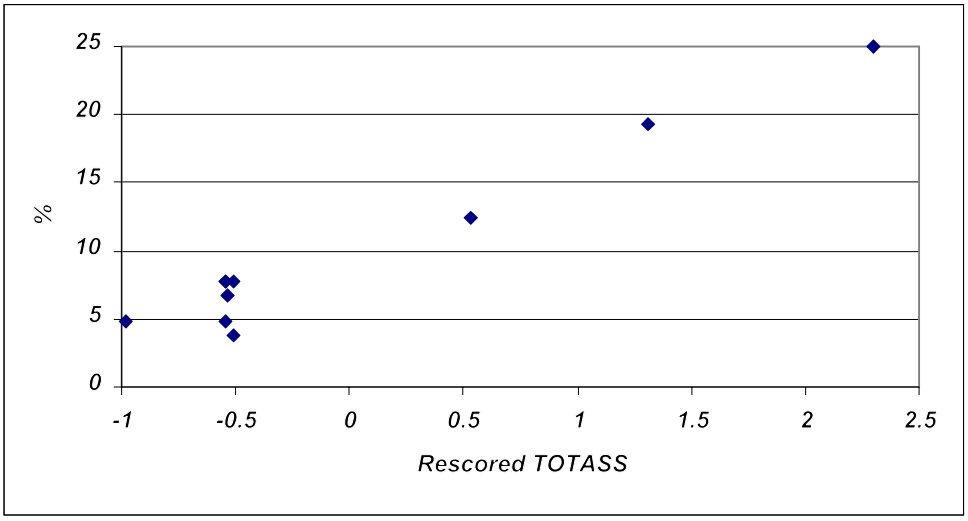


Table 2: Prediction of the MID Group				
Variable	Coefficient	Standard Error	Significance	Exp(B)
TOTASS	0.6157	0.1058	0.0000	1.8510
NETPROFM	0.3506	0.1434	0.0145	1.4199
Constant	-1.6119	0.1277	0.0000	0.1995
Source: Compustat; logistic regression				

The detail on the model includes the fact that the model involved all but 38 of the firms (some variables were eliminated because too few firms gave that data (e.g., 345 firms listed no intangible assets); mean annual wealth appreciation was not available for seven firms; 31 had no assets or net profits reported around the time of their IPO.

From this table (Table 2 above) we conclude that above average performance (MID Group membership) is associated with larger total assets and larger net profit margins – and that each (standard deviation) step up the (rescored) asset ladder results in an increase in the odds of being in the above average category (the "Exp(B) column) of 85%, while a corresponding step on the net profit scale results in an increase of the odds of 42%. (Note that the direction of the association must

be checked against the Optimal Scoring runs above.) The classification table for this regression shows that when the cut-off is adjusted so that the prediction equation puts 101 firms into the MID Group, it does so accurately in 45 (44%) of the cases. This adjustment of the posterior odds corresponds to placing the costs of misclassification other than at one-to-one. Instead, the cost of misclassifying the smaller quintile – a false negative – were put closer to four-to-one. This assignment of costs is in the spirit of recent papers advocating re-evaluation of previous simplistic cost structures (Provost, Fawcett, and Kohavi, 1998, Adams and Hand, 1999, Drummond and Holte, 2007). Overall, 413 or 79% of the 525 firms for which all of the data were available were correctly classified at this initial stage.

Table 3: Stage I – Classification Table

Predicted	Actual Others	Mid Group	Total
Others	368	56	424
MID Group	56	45	101
Total	424	101	525

Stage 2

At the second stage, the 101 firms identified by the initial logistic regression were excused from the analysis, and a new search for predictors began. The result of this search for predictors of the TOP group of IPO performers was three variables. Sales, receivable turnover (RECTURN), and the ratio of retained earnings to assets (RETE.AT) all entered the model, all did so with negative coefficients, indicating inverse relationships to the response, wealth appreciation. The details of the table below show the significance of each of the three predictors, and their impact on the probability of inclusion in the TOP quintile of mean annual wealth appreciation firms.

Table 4: Stage II – Prediction of the TOP Group

Variable	Coefficient	Standard Error	Significance	Exp(B)
SALES	-0.4169	0.1714	0.0150	0.6591
RECTURN	-0.3527	0.1544	0.0224	0.7028
RETE.AT	-0.5055	0.1313	0.0001	0.6032
Constant	-1.4973	0.1463	0.0000	0.2237
Source: Compustat, logistic regression				

The prediction equation attempts to predict 85 into the TOP group, and it is accurate in 37 (or 44%) of those predictions.

Table 5: Stage II – Classification Table				
		Actual Others	TOP Group	Total
Predicted	Others	258	53	311
	MID Group	48	37	85
Total		306	90	396

A total of 295, or 74% of the 396 firms available at this stage of the analysis, were accurately predicted at this stage.

Stage 3

The balance of the data, omitting the 101 predicted in Stage I and the 85 predicted in Stage II, were again subjected to logistic regressions with the aim of predicting membership in the lowest quintile. This LOW Group prediction equation involved two variables, current debt (DEBTCURR) and marketing expenses (SELLCOST). Perhaps intuitively, current debt is a positive indicator for the LOW group, while marketing effort is a negative indicator.

The details indicate that 270 firms remained with scores available for the analysis. Every step up of one standard deviation in DEBTCUR doubles the odds of being in the LOW Group while each similar step along the SELLCOST dimension reduces the odds by 37%.

Table 6: Stage III – Prediction of the LOW Group				
Variable	Coefficient	Standard Error	Significance	Exp(B)
DEBTCUR	0.6760	0.2389	0.0047	1.9660
SELLCOST	-0.4610	0.1705	0.0069	0.6307
Constant	-1.1580	0.1573	0.0000	0.3141
Source: Compustat; logistic regression				

The classification table shows below shows that 43 cases were predicted into the LOW Group, with 17 done so accurately, or 40% correct prediction among them. Overall, 198, or 73% of the 270 used in this analysis were correctly predicted.

Table 7: Stage III – Classification Table				
		Actual Others	LOW Group	Total
Predicted	Others	181	46	227
	LOW Group	26	17	43
Total		207	63	270

DISCUSSION AND SUMMARY

A summary of this multi-stage prediction scheme would include its ability to predict performance at three places along the continuum of mean annual wealth appreciation. At the outset, all the variables are segmented into deciles to preserve the ordinality, but destroy the exact metric of the variables. This is done for two reasons: 1) the distribution of firms along these original variables is distinctly non-normal with astronomical skewness and kurtosis numbers; and 2) it is believed that the accounting practices vary between firms and that their reported numbers bear only a tenuous relationship to the underlying reality they purport to represent. The deciles are then submitted to a re-scoring routine developed by Benezecri, Gifi, and others. This rescoring attempts to maximize a measure of monotone correlation (Kimeldorf & Sampson, 1978). It is these "quantifications," or rescorings, that are used in the balance of the analysis. The model uses different variables at each of three stages. It starts by trying to predict membership in the middle of the distribution (the sixth and seventh deciles). It is entirely possible that this non-linear relationship that is being taken advantage of is the famous "horseshoe" discussed throughout Gifi (1990). It was originally believed that if an accurate prediction could be made off the middle of the distribution, then the middle could be deleted and it would be relatively easier to predict the extremes. While this ease did not eventuate, it was possible to make respectable predictions at each stage.

The summary table for the prediction is as follows. There were at the onset, three sets of firms which were of interest: MID, TOP and LOW. The three accounted for 60% of the firms. The procedure proposed herein actually predicted 253, or 45% of the original 563 firms and was accurate in 99, or 39% of those predictions. However, when the absence of key accounting data is taken into account, the procedure actually was accurate in about three-fourths of its predictions.

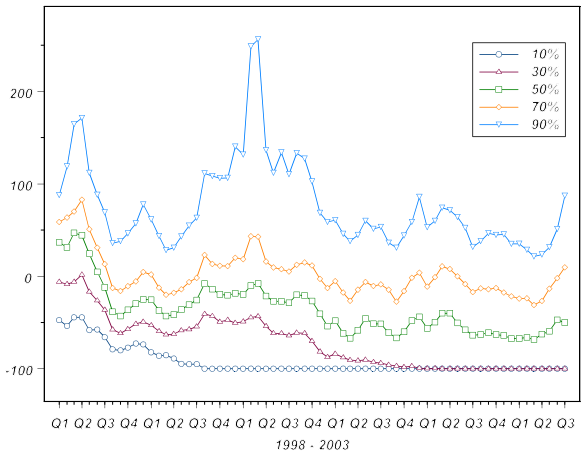
Table 8: Classification Results Summary

	STAGE I	**STAGE II**	**STAGE III**
GROUP PREDICTED	MID	TOP	LOW
Prediction Variables	*TOTASS*	*SALES*	*DEBTCURR*
	NETPROFM	*RECTURN*	*SELLCOST*
		RETE.AT	
Odds Effects	1.85	1.42	1.97
	1.42	0.70	0.63
		0.60	
Firms Available for Prediction	525	396	270
Firms in Target Group	101	90	63
Predicted Total	101	85	43
Predicted Accurately	45	37	17
Conditional Prediction Accuracy Rate	79%	74%	73%

CONFIRMATION

The original model was built on data from all IPOs issued during the calendar year 1997. Since the criterion variable of most interest was that of "annual wealth appreciation" (following Gompers and Lerner, 1999 and many others), the earliest data came from January 1998, when only 31 firms were able to show the requisite 13 months of data in order to be able to calculate the annual rate of change. By one year later all 563 firms were showing wealth appreciations. By the end of the model period, January 1998 up to February 2001, virtually all of the 1990s "bubble" had evaporated, and the "interim" period from March 2001 up to July 2003 saw a relatively subdued market with most of the 1997 class of IPO stocks taking money from their investors. (Table below.) As the table demonstrates, only the top quintile was performing at annual rates above zero (returning one dollar in market value plus accumulated and reinvested dividends for one dollar invested at the beginning of the 12 month period). Over 40% had essentially disappeared from the markets, either through failure, merger, purchase, or any of the other exit routes from the markets. This exodus posed a difficult challenge for the model to perform well.

Figure 7: Mean Performance by Quintiles



The present confirmation study involved inspection of data from Standard & Poor's Research Insight, and included all but 24 of the original IPOs, which showed stock price and other company information up through July, 2003. It might be expected that the missing companies would be from the lower performing groups. In fact, though, seven were from the TOP group, three from the MID group, and nine were from the LOW group.

The additional data provided nearly 1.5 years (17 months, the so-called "interim" period) more data than the earlier analysis, which had a cut-off date of February, 2001.

A more serious concern was that many of the companies were missing one or more months of wealth appreciation data. In fact, 164 or 30.4% were missing at least 20 of the months (the maximum missing was 29 months). Again, however, there was no indication that the absence of data was related to the performance of the company's stock. The percentages of those missing 20 data points or more was 28%, 35%, and 26% in the TOP, MID and LOW groups, respectively (these groups are those established in terms of their mean performance over the original three years). And, the same missing class represented 25%, 31%, and 27% of those firms predicted to be in the TOP, MID and LOW groups, respectively.

The missing data, then, was not seen as a differentially distorting factor. All missing data (from 1999 on) were replaced by the value "-100," representing the depressing fact that for most of these firms their investors would have lost all of their investment had they held onto their stock for the full 12 months. Discriminating between the various shades of disappearances will be the subject of a future analysis. For the purposes of this confirmation analysis, however, any distortions due to this oversimplification will simply be absorbed into the error estimation for the model.

The Model Tested. Two tests formed the basis of the confirmation of the model: 1) study of the mean shifts of the predicted groups since they were formed in mid-2001; and 2) study of the composition churn in the predicted groups since mid-2001.

Questions to be answered are 1) are the predicted groups performing at their (relative) predicted levels over the full five-year period? 2) does the model still predict full five-year performance better than chance?

As might be expected from the overall patterns, the predicted groups performed much differently in the "interim" period than they had during the original three years. In the original period the TOP group demonstrated a mean annual growth of about 50%, while the MID group (actually the sixth and seventh percentiles) was nearly even and the LOW group depreciated at about a third per year. All three groups declined during the interim period, and they did not all decline at the same rate (p-value = .03 for a simple ANOVA including only the data from the three groups). The TOP group actually declined the most, losing more than 50 percentage points from its mean Wealth Appreciation level during the original period. The MID group dropped the least, but still performed at a level about 30 percentage points below its pre-2001 level. The LOW group lost performance at a rate mid-way between TOP and MID.

However, despite the distortions caused by the differential interim performances, the five-year means retained their original relative standings. The TOP group lost all investor advantage they had had in the earlier period to return essentially what was invested, while the MID group slid to a mean loss level of about a sixth of what was invested 12 months earlier. The LOW group maintained both its relative position at the bottom, and its absolute level of draining about one-third of its investors' money each year. The statistical test of these five-year performances was very significant (p-value under .0001).

Table 9: Mean Shifts in Wealth Appreciation				
	1997 to Feb 2001	Interim to 2003	1997 to July 2003	Number of Firms
Predicted Group				
TOP	51.20	-51.44	2.73	75
MID	-5.26	-31.61	-17.96	98
LOW	-33.87	-44.77	-38.82	43
Other	-5.35	-31.44	-17.57	324
Total	0.25	-35.31	-16.52	540
ANOVA tests of the means from the three prediction groups alone have F ratios and p-values of 14.114, 0.00000175; 3.542, 0.0306; 7.814, 0.000531, respectively.				

Finally, note that the range between the means of the TOP and LOW predictions groups dropped from about 85 percentage points to about 41.5 percentage points. This is a very striking reduction and is no doubt at least partially due to some sort of "regression to the mean" effect. However, the entire market for IPOs went through a compression relative to the earlier period. For example if one looks at the "interdecile range" (difference between the tenth and ninetieth percentiles), its value for the original period was about 148 percentage points. The value for the interim period was about 135 points while the overall five-year interdecile range was only 118 percentage points (not shown).

Further, a comparison of the performance groups' composition based on all five years of wealth appreciation shows the same pattern as the means above. While the original model correctly predicted 43.2% of the training set (99 out of 229 predicted into one of the groups) into the three quintiles, the quintiles based on the full five years' data were correctly predicted 29.6% of the time (64 of the 216 firms in the TOP, MID and LOW quintiles were accurately predicted).

Under the random model, in which only 20% of the quintile predictions should be accurate, a chi square goodness-of-fit test of the original model had a test statistic value of 88.6 with a p-value starting with 20 zeros, indicating a very high significance. The test of the updated model had a chi square of 12.5 with a p-value of 0.0004.

Table 10: Prediction Accuracy – Wealth Appreciation Quintiles		
	1997 to Feb 2001	1997 to July 2003
	Correct	TotalCorrectTotal
TOP	37	852375
MID	45	1012798
LOW	17	431443
Total	43.2%	100%29.6%100%

DISCUSSION

The purpose of this empirical analysis was to evaluate a model created in mid-2001 which has as its objective the prediction of annual wealth appreciation performances over a three-year period (January 1998 to February 2001) of 563 IPOS issued in 1997. Since creation of the model, the markets have suffered through a prolonged period of poor returns to its investors. In fact, the data do suggest that forecasting during the pre-2001 period may have been more likely to be successful than the "interim" period since. The great bulge in the performance of the TOP group (top quintile) had disappeared in early 2001. However, the strength of the stocks during 1998-2001 was enough that the five-year (actually January 1998 to July 2003) means still reflected that earlier performance. Perhaps it is a further mark of the potency of the model that it was able to weather these extremes.

Like all models there are more questions than answers. While this empirical test may have contributed to the question about whether any modeling might be effective, we still are concerned about generalizing it. There are structural issues about using the techniques in other times and for other types of equities. There are substantive questions about the cause and effect – the process by which the variables utilized in the prediction scheme materialize and lead to the results in the markets.

Future research is very much needed in several areas. To start with, it would be of great interest to see if the non-metric, "deconstructive" methods used to develop this model will be similarly successful using the full five-years' data. And, if successful, is the five-year model similar to the earlier one? Does the new model "find" the same groups? Does it employ the same, or related, variables? (These variables are, to put it mildly, inter-correlated. Each of the 26 predictors considered relates to one of only a handful of underlying factors. While an untangling of the variable intercorrelations might shed light on which of the variables is most or least effective, the overall model strength should not be affected.

Use of alternative classification algorithms to the logistic regression techniques is also worth considering. Some recent research has found the new boosting procedures useful in predicting corporate failures (Cortes et al., 2007). On the other hand, it should be noted that the inherent strengths of older techniques like logistic regression have been shown both in a study of successful companies very similar to this one (Johnson and Soenen, 2003), and in more basic research coming out of the statistical and machine learning communities (Holte, 1993, Lim et al., 2000).

More work needs to be done on the "missing" firms. Not all of the missing firms suffered catastrophic declines. Perhaps more artful estimates of the transformed entities derived from the original 563 IPOs will provide more insight into the model's accuracy.

Not only the missing group, now accounting for nearly half of the 1997 IPOs, are of interest. It would be instructive to follow-up on the TOP group. How many of those in the TOP group are still up there? How does the turmoil in TOP membership relate to the prediction model? What characterizes the TOP firms which maintained versus those that slid?

REFERENCES

Adams, N. M. & Hand, D. J. (1999). Comparing classifiers when the misallocation costs are uncertain. *Pattern Recognition*, 32, 1139-1147.

Agresti, A., C. Chuang & A. Kezouh. (1987). Order-restricted score parameters in association models for contingency tables. *Journal of the American Statistical Association*, 82, 619-623.

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

Bardos, Mireille (2000). Detection of company failure and global risk forecasting, in *Data Analysis, Classification and Related Methods*, Kiers, Henk A.L., Rasson, Jean-Paul, Groenen, Patrick J. F. and Schader, Martin (Eds.). Berlin, Springer.

Beaver, W. (1967). Financial ratios as predictors of failure, *Empirical Research in Accounting, Selected Studies, Supplement to Journal of Accounting Research*, 5, 71-111.

Bridgman, P. (1927). *The logic of modern physics*, New York, Macmillan.

Chatfield, C. (1995). Model uncertainty, data mining and statistical inference with comments, *Journal of the Royal Statistical Society, Series A (Statistics in Society)*. 158,419-466.

Cortes, Alfaro E., Gamez Martinez, M. & Garcia Rubio, N. (2007). Multiclass corporate failure prediction by Adaboost.M1. *International Advances in Economic Research*, 13, 301-312.

Cox, D. & D. Oakes (1984). *Analysis of survival data*, London, Chapman & Hall.

de Leeuw, J. (1993). Some generalizations of correspondence analysis. Retrieved April 15, 2003 from http://citeseer.nj.nec.com/deleeuw93some.html.

Drummond, C. & Holte, R. R. (2007) Cost curves: An improved method for visualizing classifier performance. *Machine Learning*, 65, 95-130.

Edmister, R. (1972). An empirical test of financial ratio analysis to small business failure prediction. *Journal of Financial and Quantitative Analysis*, 7, 1477-1493.

Fisher, R. (1940). The precision of discriminant functions. *Annals of Eugenics*, 10, 422-429.

Fisher, F. (1986). Statisticians, econometricians, and adversary proceedings. *The American Statistician*, 81, 277-286.

Gifi, A. (1990). *Nonlinear multivariate analysis.* Chichester, John Wiley & Sons.

Gilula, Z., A. Krieger. and Y. Ritov (1988). Ordinal association in contingency tables: some interpretive aspects. *Journal of the American Statistician*, 83, 540-545.

Gompers, P. (1995). Optimal investment, monitoring, and the staging of venture capital. *Journal of Finance*, 50, 1461-1490.

Gompers, P. and J. Lerner (1999). *The venture capital cycle*. Cambridge,Mass., The MIT Press.

Hand, D. (1996). Statistics and the theory of measurement (with comments). *Journal of the Royal Statistical Society, Series A*, 159,445-492.

Hirschfeld, H. (1935). A connection between correlation and contingency, *Proceedings of the Cambridge Philosophical Society*, 31,520-524.

Holte, R. C. (1993). Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11, 63-90.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 417-441, 498-520.

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28, 321-377.

Johnson, R. & Soenen, L. (2003). Indicators of successful companies. *European Management Journal*, 21, 364-369.

Kimeldorf, G. & A. Sampson (1978). Monotonic dependence. *Annals of Statistics*, 6, 895-903.

Krantz, D., R. Luce, P. Suppes & P. Tversky (1971). *Foundations of measurement, vol. 1, Additive and polynomial representations*. New York, Academic Press.

Lakonishok, J., J. Shleifer & R. Vishny (1994). Contrarian investment, extrapolation and risk. *Journal of Finance*, 49, 1541-1578.

Lim, T-S., Loh, W-Y. & Shih, Y-S. ((2000). A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine Learning*, 40, 2003-228.

Luce, R., D. Krantz, P. Suppes & A. Tversky (1990). *Foundations of measurement, vol. 3, Representation, axiomatization, and invariance*. San Diego: Academic Press.

McCullagh, P. (1980). Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society, Series B*, 42, 109-42.

Meulman, J. (1992). The integration of multidimensional scaling and multivariate analysis with optimal transformations. *Psychometrika*, 57, 539-565.

Miller, J. M. (2003). *Venture capital, entrepreneurship, and long-run performance prediction:An application of data mining*. Unpublished doctoral dissertation, Rice University.

Press, S. & S. Wilson (1978). Choosing between logistic regression and discrimination analysis. *Journal of the American Statistical Association*, 79, 699-705.

Provost, F., Fawcett, T. & Kohave, R. (1998). The case against accuracy estimation for comparing induction algorithms. In: *Proceedings of the Fifteenth International Conference on Machine Learning*, 43-48.

Ritov, Y. & Z. Gilula (1991). The order-restricted RC model for ordered contingency tables: Estimation and testing for fit. *Annals of Statistics*, 19, 2090-2101.

Ritov, Y. & Z. Gilula (1993). Analysis of contingency tables by correspondence models subject to order constraints. *Journal of the American Statistical Association*, 88, 1380-1387.

Stevens, S. (1951). Measurement, statistics, and psychophysics, In Stevens, S. (Ed.), *Handbook of Experimental Psychology*, New York: John Wiley & Sons.

Suppes, P., D. Krantz, R. Luce & A. Tversky (1989). *Foundations of measurement, vol. 2, Geometrical, threshold, and probabilistic representations*. San Diego: Academic Press.

Tukey, J. (1977). *Exploratory data analysis.* Reading, Mass.: Addison-Wesley.

Tukey, J. (1980). We need both exploratory and confirmatory. *The American Statistician*, 34, 23-25.

Velleman, P. & L. Wilkinson (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, 47, 65-72.

Zopoundis, C. & A. Dimitras (1998). *Multicriteria Decision Aid for the Prediction of Business Failure*. Boston: Kluwer Academic Publishers.

DO SHORT SELLERS ANTICIPATE LARGE STOCK PRICE CHANGES?

Roger J. Best, University of Central Missouri
Ronald W. Best, University of West Georgia
Jose Mercado-Mendez, University of Central Missouri

ABSTRACT

It is commonly believed that short sellers must possess better information than the average trader in order to overcome the costs associated with short sale restrictions. Previous studies provide evidence which is consistent with, but insufficient to prove, informed trading. Because the role of short sellers in the stock market has come under greater scrutiny by regulators, understanding the nature of the information used by these traders is crucial in determining whether there is need for additional restrictions on short sellers.

We examine short interest around large, one-day stock returns. Our methodology allows us to overcome biases that may exist in previous studies. We find significant and robust increases in short interest immediately prior to large negative returns caused by a limited subset of corporate events, but no statistically significant increases in short interest for firms with negative returns caused by commonly occurring events. Surprisingly, the level of short interest is higher for firms experiencing large, positive returns, and we find no significant declines in short interest preceding these positive returns. Thus, we find weak evidence of informed trading by short sellers, but must conclude most short selling activity appears to result from speculation or hedging and not informed trading.

INTRODUCTION

We examine whether short sellers anticipate large changes in stock prices. That short sellers must possess superior information has been generally accepted since Diamond and Verrecchia (1988) and Fabozzi and Modigliani (1992). These authors show that in the presence of short sale restrictions, short sellers must have an information comparative advantage in order to overcome the costs of these restrictions. Documented stock market reactions appear to be consistent with this notion. Conrad (1986) and Senchack and Starks (1993) find that there are significantly negative stock price changes around announcements of large levels of and changes in short interest. Further, Desai et al. (2002) show that stocks listed on NASDAQ with relatively large levels of short interest

experience longer-term abnormal returns. That short sellers possess more precise information is given further credence by Kadiyala and Vetsuypens (2002), who use short interest as a measure of information near stock splits. In general, these authors find that the level of short interest changes around certain splits, leading the authors to conclude that some stock splits provide information to market participants.

The question of whether short sellers do possess superior information, however, is not yet resolved. While negative returns following short sale announcements are a necessary condition to show short sellers are better informed, the negative returns alone are not sufficient to *prove* the superior information hypothesis. The use of ex post stock market reactions such as those in Senchack and Starks (1993) may lead to erroneous conclusions. For example, Chan (2003) documents that market participants tend to react incorrectly to news, particularly negative news. Further, it is now well known that investors suffer from numerous cognitive biases (see Hirshleifer 2001 for a complete description). It is plausible, therefore, that returns surrounding short interest announcements offer an incomplete picture of the news contained in short interest–that is, market participants may inaccurately assume that short sellers possess a better information set and adjust stock prices based simply on the observation of an increase in short sales. Indeed, recent papers suggest that short sellers may be no better informed regarding the prospects of the average firm.

Mercado-Mendez, Best and Best (2006) find that, in general, short sellers do not increase short positions in the month prior to large negative earnings surprises. Further, the levels of short interest in firms with large negative earnings surprises are similar to those for firms that have large positive earnings surprises. Closer inspection reveals that only for a subset of firms with extreme negative earnings surprise does short interest exhibit patterns consistent with informed trading. Christophe, Ferri and Angel (2004), using daily short interest data for NASDAQ-listed firms, find increased short selling prior to earnings announcements that generate negative stock price changes. However, they also find similar activity for some firms that have positive returns around earnings announcements. These findings question whether the average short seller possesses better information than other traders.

To resolve this issue, we examine the pattern of short interest around large stock price changes. Presumably, short sellers would be most interested in shorting stocks that will have a near term price decline and would avoid stocks that will have a near term price increase. We also categorize the large returns by the events that (apparently) led to those returns. By focusing first on large price changes (as opposed to earnings announcements as in previous studies) and conditioning on underlying causes of the price change, we can determine whether short sellers are better able to anticipate some corporate events. Further, we use the relative level of short interest, unlike Christophe et al. (2004) who examine short sale transactions on a daily basis but cannot determine whether the level of short interest is changing. Using the level of short interest (as opposed to individual short sale transactions), we can infer whether the typical short seller engages in informed or speculative trading.

DATA AND METHODOLOGY

We begin our sample collection by including firms traded only on the NASDAQ system. We use these firms because of the readily available short interest data provided electronically by NASDAQ during this time period. Given the similarity of short sale restrictions between NASDAQ and the NYSE, we presume our findings would generalize to other exchanges. We reduce the universe of firms further by eliminating firms that have average daily volume of less than 100,000 shares. This eliminates a group of firms that short sellers would have greater difficulty in finding shares to borrow and increases the likelihood of complete data availability. From this remaining universe of stocks, we use the CRSP database to identify the firm with the most extreme positive stock return and the firm with the most extreme negative stock return on each trading day from the 15th to the 21st of each month over the period April 1996 - July 2002. We use these dates because short interest data is compiled on the 15th of each month and reported on (approximately) the 22nd of that month. This allows us to more closely align the event date (extreme return) with the reported short sales data without contaminating observed returns with the reporting of short interest.

We exclude observations for several reasons. First, if a firm appears in the sample less than four calendar months from a previous occurrence, we assume that this subsequent observation is contaminated by previous information and delete the second occurrence. Second, firms must have financial data available on Standard and Poor's *Research Insight* database so that we may collect market value of equity and book-to-market value. Third, stock prices and returns must be available on CRSP for six months prior to the extreme return. Finally, the number of shares sold short for the months surrounding the large return must be available from NASDAQ. In particular, we require the level of short sales to be available from four months prior to two months after the date on which the extreme return occurs. This leaves 220 firms with extreme negative returns and 199 firms with extreme positive returns for our sample.

To ascertain whether short sellers anticipate large stock price changes, we utilize several tests. First, we examine the level of short interest in the three months prior to, month of, and the two months immediately following the large stock price change. We report short sales divided by most recently reported shares outstanding (from CRSP) instead of the commonly used average trading volume standardization. We standardize by shares outstanding to eliminate the possibility that observed changes in the short interest ratio result from changes in trading volume instead of changes in the amount of shares sold short, particularly important given the long time period over which we measure short interest.

Next, we calculate the average change in our short interest measure from month to month and use t-tests to determine whether observed changes in the short interest ratio are statistically significant. If short sellers are better informed, we expect to find significant increases in short interest prior to extreme negative returns. The prediction for firms with extreme positive returns is more ambiguous. If short sellers are already avoiding stocks that will experience large positive returns, we would likely find little or no change in short interest around the return date. Otherwise,

if significant short interest exists, large declines in short interest prior to positive returns would be consistent with informed selling. Thus, if short sellers trade on better information, we should find short interest immediately before the large return to be significantly higher for stocks that experience extreme negative (one-day) returns than for stocks with extreme positive (one-day) returns.

After examining changes in short interest for the full sample, we ascertain whether short sellers are better able to anticipate certain events by identifying the cause of the substantial change in stock price. To identify the return driver, we search the Lexis-Nexis database for news wire reports around the date of the extreme return. We classify firms based on common events and examine the pattern of short interest in the months around the event date based on these classifications.

EMPIRICAL RESULTS

To determine whether there is industry clustering, we categorize our sample by SIC Code. The number of firms that fall within a particular 1-digit SIC code is included in Table 1 by positive and negative return companies. As shown, there are several SIC groupings where clustering appears to occur. In particular, over half of our sample observations are from the 3-digit and 7-digit SIC code groupings. These groupings include firms generally regarded as "high-tech" firms and traditional manufacturing firms. The relative proportions of all firms, however, are generally consistent with those in Mercado-Mendez et al (2006), who report that their results are not influenced by industry effects. Further, we note that the proportions of firms in each SIC grouping are remarkably similar for both positive and negative returns.

Table1: Sample firms by SIC Code

SIC Code Range	Negative Returns	Percentage of Negative	Positive Returns	Percentage of Positive	All Firms	Percentage of All Firms
1000 - 1999	2	0.9	2	1.0	4	1.0
2000 - 2999	19	8.6	18	9.0	37	8.8
3000 - 3999	92	41.8	66	33.2	158	37.7
4000 - 4999	16	7.3	14	7.0	30	7.2
5000 - 5999	14	6.4	10	5.0	24	5.7
6000 - 6999	1	0.5	3	1.5	4	1.0
7000 - 7999	66	30.0	72	36.3	138	32.9
8000 - 8999	10	4.5	14	7.0	24	5.7
Total	220	100%	199	100%	419	100%

The sample includes firms that experience large positive or negative one-day returns on any day from the 15th to the 21st day of the month over the period April 1996-July 2002.

In Table 2, we provide descriptive information on the sample based on whether the one-day return is negative or positive. The average (median) return for the negative return group is -21.2% (-17.6%). The returns range from -62.3% to -4.5%. Given these extreme returns, we assume short selling in advance of these one-day declines would generate economically significant profits. Not surprising given our selection of NASDAQ-only firms, the average negative-return firm in our sample has a market value of equity of only $2.13 billion. The distribution of market values is skewed as the median firm has a market value of approximately $1 billion. The average (median) book-to-market ratio for these firms is 0.212 (0.16) while the average (median) shares short (as a percentage of shares outstanding) is 5.9% (3.8%) in the month of the return date. There is wide variation in short interest in our sample, with a low of 0.2% and a high of over 33% during this month.

Table 2: Descriptive Information

Panel A: Firms with Negative Returns				
	Mean	Median	Minimum	Maximum
Return	-21.2%	-17.6%	-62.3%	-4.5%
Market Value (in millions)	2,130.3	1,042.1	38.2	26,490.5
Book-to-Market	0.212	0.160	-0.382	0.979
SI	5.9%	3.8%	0.2%	33.7%
Panel B: Firms with Positive Returns				
Return	24.5%	20.5%	6.2%	110.4%
Market Value (in millions)	1,499.7	659.6	17.4	38,955.9
Book-to-Market	0.278	0.201	-0.441	1.946
SI	6.8%	4.2%	0.1%	41.6%

Return is the raw one-day extreme return identified from the CRSP database. Market Value is the market value of common equity in millions of dollars, measured approximately two months prior to the return date. Book-to-market is the book to market ratio measured approximately six months prior to the return date and SI is short interest from the month of the return calculated as cumulative shares short divided by shares outstanding during the same month (reported as a percentage).

For firms that experience positive returns, the mean (median) one-day return is 24.5% (20.5%) with a range of 6.2% to 110.4%. The average size of these firms is approximately $1.5 billion, which, statistically, is significantly smaller than the market value of equity for the negative-return firms. Further, the average book-to-market ratio of 0.278 is significantly greater than the book-to-market ratio of the negative-return firms. Interestingly, it appears that the typical firm in the positive-return group has a higher short interest than the typical negative-return firm immediately before the large return. The average (median) short interest for the positive-return

firms is 6.8% (4.2%). Thus, the sample firms with large positive returns, in general, are smaller, have lower valuations, and have higher levels of short interest than the sample firms with large negative returns.

Table 3: Average Short Interest Relative to the Return Date

	SI-3	SI-2	SI-1	SI0	SI+1	SI+2
Negative Return Firms	4.9% (3.0%)	5.2% (3.4%)	5.6% (3.7%)	5.9% (3.8%)	5.9% (4.0%)	6.1% (4.0%)
Positive Return Firms	6.2% (3.8%)	6.1% (3.7%)	6.7% (4.2%)	6.8% (4.2%)	7.5% (4.6%)	7.2% (4.5%)
t-test for differences (Negative - Positive)	-1.982[b]	-1.492	-1.668[c]	-1.354	-2.284[b]	-1.667[c]

Short Interest is the average shares sold short as a percentage of shares outstanding. Median values appear in parentheses. SI0 represents short interest in the month of the extreme return, SI-3 represents short interest three months prior to the extreme return, and SI+2 indicates short interest two months after the extreme return. Short shares are monthly values as reported by NASDAQ. [b] Significant at 5%, [c] Significant at 10%

We next examine the level of short interest in the months surrounding the extreme return and statistically compare these levels across return-type. Specifically, we examine the average short interest in the three months prior, month of and two months following the extreme return. Based on our methodology, if short sellers are better informed, we expect changes in short interest to occur in the month of the large return, and possibly in the month prior (if borrowing constraints exist, short sellers may have to act upon the information over longer horizons than would otherwise be optimal). The results of this analysis appear in Table 3. For the negative return firms, both the average and median short interest increases in the months leading up to and including the extreme return, which is consistent with informed trading. Interestingly, however, we observe a similar pattern for firms that experience a positive return. Even more surprising, the average short interest for the positive return firms is *greater* than the average short interest of negative return firms in every month. This difference is statistically significant in four of the six months (-3, -1, +1, and +2) that we examine. While the post-event differences might be explained by reassessments of a near-term price decline for positive return firms or a reduction in short interest for negative return firms, it is difficult for us to reconcile the pre-event differences in short interest with the notion of informed trading.

To further investigate these findings, we calculate the change in average short interest from month-to-month to verify whether the increases we observe in Table 3 represent statistically significant changes. We report the monthly (average) change in Table 4 by return type. For the months leading up to and including the large negative returns, the change in average short interest is statistically significant. Thus, it appears that short sellers, at least to some degree, are anticipating that these firms will experience large price declines. We note, however, that for firms with large

positive returns, average short interest increases significantly in the month prior to and remains statistically the same in the month of the large return. Thus, we are unsure why informed short selling might exist for companies with "bad" news when there is an apparent inability of short sellers (on average) to anticipate large positive stock price changes.

Table 4: Change in Percentage Short Interest Surrounding the Return Date

	SI-2	SI-1	SI0	SI+1	SI+2
Negative Return Firms	0.284 (1.972)[b]	0.421 (3.287)[a]	0.290 (1.841)[c]	-0.007 (-0.050)	0.222 (2.287)[b]
Positive Return Firms	-0.173 (-0.699)	0.657 (2.103)[b]	0.059 (0.195)	0.685 (2.988)[a]	-0.233 (-1.096)

Short Interest is the average short shares as a percentage of shares outstanding. The change in short interest is calculated as current month percentage short interest minus prior month percentage short interest. t-statistics, testing whether the observed change is significantly different from zero, appear in parenthesis. Short shares are monthly values as reported by NASDAQ.

[a] Significant at 1%, [b] Significant at 5%, [c] Significant at 10%

To provide insight into this puzzle, we examine changes in short interest based on the events which appear to have caused the large returns of the firms in our sample. From Lexis-Nexis, we are able to find news reports for 146 (of 199) positive-return firms and 154 (of 220) negative return firms. Based on these news reports, we classify firms into one of four news-event categories: (1) Earnings-Related, (2) Strategic Change, (3) Product-Related, and (4) Miscellaneous. Earnings-related announcements include actual earnings announcements, earnings guidance or warnings, analysts' reports, and sales guidance. Strategic change announcements include mergers, acquisitions, strategic alliances and management changes. Product-related news involves progress reports on clinical trials and new product contracts. The miscellaneous category includes all other events. For our sample, examples of miscellaneous events are financing announcements, lawsuits, insider trades, restructuring, and stock splits. Table 5 lists the number of firms and average one-day return for each news group. As shown, the highest percentage of positive (30.7%) and negative (50.9%) return firms falls into the earnings-related announcements category. The miscellaneous category, however, has the largest (absolute) average return across both return types with average returns of 37.5% and -24.5%.

In Table 6, we report the level of short interest in the months including and surrounding the extreme return based on the news category and for firms that have no (identifiable) corresponding news announcement. For the negative return firms in Panel A, an analysis of variance F-test indicates that there are significant differences in the level of short interest across categories. The larger levels of short interest appear to be concentrated in the product-related and miscellaneous categories. For positive return firms in Panel B, the levels of short interest across categories are not

statistically different. Interestingly, however, the largest level of short interest does appear concentrated in the "no news" category.

Table 5: Events Leading to Extreme Return

	Positive Returns			Negative Returns		
Event	Num. Obs.	% of Events	% Mean Return	Num. Obs.	% of Events	% Mean Return
0. No Accompanying News Report	53	26.6	20.2	66	30.0	-16.3
1. Earnings-Related	61	30.7	23.1	112	50.9	-24.3
2. Strategic Change	32	16.1	28.5	21	9.5	-17.3
3. Product-Related News	40	20.1	25.0	12	5.5	-23.5
4. Miscellaneous Events	13	6.5	37.5	9	4.1	-24.5
All Events	199	100.0	24.5	220	100.0	-21.2

Event 1 includes earnings announcements, restatements, equity issuances, and analyst rating changes. Event 2 includes mergers, acquisitions, alliances, and management changes. Event 3 includes all announcements related to products (such as clinical trial results for pharmaceutical companies), and Event 4 includes all other announcements (including lawsuits and stock splits).

We next examine the monthly change in short interest based on the news categories. These results appear in Table 7. As shown in Panel A, we observe significant increases in average short interest for various months and news events for firms experiencing large negative returns. For the "no news" group of firms, short interest increases by a statistically significant 0.486 percentage points in the month prior to a large negative return and by an additional 0.340 percentage points in the month of the return. The earnings-related news group has a significant increase of 0.324 percentage points in the month prior to the negative return, while the strategic change announcements and miscellaneous announcements are associated with significant increases of 0.879 percentage points and 2.820 percentage points respectively in the month of the large negative return. In Panel B, which includes firms with positive returns, we show that there are no significant changes in the level of short interest, regardless of category, for two months prior, one month prior and the month of the large return. As we show in Table 6, several of the news categories for positive return firms have higher short interest than for the negative return firms. Thus, while some of the changes in Panel A of Table 7 are consistent with short sellers possessing superior information, the lack of significant reductions in short interest from Panel B leads us to question the true nature of the information set used by the typical short seller.

To help answer this question, we utilize regression analysis so that we may examine the joint impact of a number of factors. These results appear in Table 8. For this analysis, we use only the negative return firms and estimate six separate regression equations. While we remain puzzled by the short interest associated with positive-return firms, we use only negative return firms in the

regression to determine whether we can isolate any informed trading. In the first three equations, our dependent variable is the level of short interest (as a percentage of shares outstanding) in the month of the extreme return. In the remaining three equations, the dependent variable is the change in the percentage of shares sold short over the three-month period that ends with the month of the large return.

Table 6: Average Short Interest Relative to the Return Date by Event

Panel A: Negative Return Firms						
	SI-3	SI-2	SI-1	SI0	SI+1	SI+2
0. No News Report	4.2%	4.5	5.0	5.3	5.4	5.5
1. Earnings-Related	5.0%	5.0	5.3	5.5	5.4	5.5
2. Strategic Change	4.9%	5.4	5.8	6.7	6.4	7.0
3. Product-Related	8.2%	9.3	8.8	7.1	7.4	8.0
4. Miscellaneous	4.7%	6.8	9.3	12.2	13.0	13.8
ANOVA F-Test	1.401	**2.199**[c]	**2.106**[c]	**3.129**[b]	**4.058**[a]	**4.592**[a]
Panel B: Positive Return Firms						
	SI-3	SI-2	SI-1	SI0	SI+1	SI+2
0. No News Report	9.2%	8.1	8.3	8.3	9.1	8.9
1. Earnings-Related	5.4%	5.6	6.8	6.7	7.4	7.0
2. Strategic Change	4.0%	4.1	4.5	4.8	5.1	4.6
3. Product-Related	5.8%	5.9	7.0	7.1	7.9	8.2
4. Miscellaneous	4.8%	4.9	4.5	4.6	5.0	4.8
ANOVA F-Test	**2.810**[b]	**2.433**[b]	1.655	1.599	1.683	**2.250**[c]

Short Interest is the average shares sold short as a percentage of shares outstanding. SI0 represents short interest in the month of the extreme return, SI-3 represents short interest three months prior to the extreme return, and SI+2 indicates short interest two months after the extreme return. Short shares are monthly values as reported by NASDAQ. Events are determined based on news wire reports or articles appearing in print news services. Event types are as described in Table 5. [a] Significant at 1%, [b] Significant at 5%, [c] Significant at 10%

In the first regression (labeled regression 1 in Table 8), our only independent variable is the extreme (negative) return. Our presumption is that if informed short selling exists, short sellers would tend to short more heavily those stocks that are expected to have the largest (percentage) price decline. As the regression estimates indicate, however, there is no statistical relationship between the magnitude of the one-day return and the level of short interest in the month of that return. In our second regression, along with the extreme one-day return, we include a number of factors that may be correlated with short interest. These other factors include: (1) The book-to-market ratio (BM), shown by Reed (2003) to be related to short interest; (2) The natural log of the number of shares

outstanding, which we presume could be related to short interest because short sellers are better able to trade on private information when a greater number of shares are available for borrowing; (3) The compounded stock return for the six months prior to the news event date (Prior Return), which allows us to control for unusual "run up" (or run-down) effects; and (4) The standard deviation of the Prior Return, which may indicate the degree of information asymmetry associated with a particular company. For firms with higher levels of information asymmetry, the revelation of news typically leads to greater stock price responses, implying a greater profit potential from shorting the stocks of these firms.

Table 7: Change in Percentage Short Interest Surrounding the Return Date by Event

Panel A: Negative Return Firms			
	SI-2	SI-1	SI0
0. No News Report	0.265 (1.222)	**0.486** (2.626)[b]	**0.340** (1.985)[c]
1. Earnings-Related	0.008 (0.053)	**0.324** (1.906)[c]	0.157 (0.804)
2. Strategic Change	0.535 (0.965)	0.392 (1.429)	**0.879** (2.197)[b]
3. Product-Related	1.133 (1.432)	-0.545 (-0.877)	-1.676 (-1.074)
4. Miscellaneous	2.135 (1.188)	2.520 (1.741)	**2.820** (2.480)[c]
Panel B: Positive Return Firms			
	SI-2	SI-1	SI0
0. No News Report	-1.033 (-1.243)	0.198 (0.649)	0.027 (0.094)
1. Earnings-Related	0.170 (0.732)	1.149 (1.387)	-0.080 (-0.106)
2. Strategic Change	0.148 (0.669)	0.324 (0.623)	0.363 (0.671)
3. Product-Related	0.087 (0.239)	1.145 (1.670)	0.061 (0.077)
4. Miscellaneous	0.126 (0.445)	-0.454 (-1.072)	0.087 (0.693)

Short Interest is the average short shares as a percentage of shares outstanding. The change in short interest is calculated as current month percentage short interest minus prior month percentage short interest. t-statistics, testing whether the observed change is significantly different from zero, appear in parenthesis. Short shares are monthly values as reported by NASDAQ. Events are determined based on news wire reports or articles and are as described in Table 5.

[a] Significant at 1%, [b] Significant at 5%, [c] Significant at 10%

As the estimates from regression 2 indicate, the magnitude of the return is again not significantly related to the level of short interest. Unlike Reed (2003), we also find no relation between the book-to-market ratio and the level of short interest in the month of the large return. We do, however, find a significant negative relationship between the level of short interest and both (the natural log of) shares outstanding and the prior return. While we might suspect that with fewer shares outstanding short sellers would have greater difficulty in borrowing shares, it appears that short interest is actually relatively higher for these firms. In results not elsewhere reported, we find a significant positive correlation between market value of equity and shares outstanding for the sample of firms with negative returns. Thus, it appears that short sellers may be targeting smaller firms. Further, the more negative the prior stock performance, the greater the level of short interest, which implies that short sellers expect continued deterioration of stock price for stocks which have performed poorly, or, alternatively, short sellers appear to avoid stocks that have had recent increases in share price.

In the interest of parsimony, when we estimate regression 3, we exclude variables that have insignificant coefficients in regression 2. Thus, the independent variables in our third regression consist of the natural log of shares outstanding and the six-month prior return. We also add dummy variables indicating the different events that we have identified. D0 equals one if there is no news report accompanying the large return, D1 equals one if the large return results from earnings-related news, D2 equals one if the announcement indicates a strategic change, and D4 equals one for miscellaneous news reports. As shown, the coefficients for shares outstanding and the prior stock return remain negative and significant in the presence of the dummy variables, implying that short sellers trade on publicly available information. We also find that coefficient for D4 is positive and significant. Thus, it appears that short interest is higher (on average) for firms experiencing irregular, uncommon negative events after controlling for prior returns and borrowing restrictions. None of the other dummy variables have statistically significant coefficients.

We next explore whether short sellers tend to increase short positions in the months leading up to the large negative return. For regressions 4, 5, and 6, we use the three-month change in the proportion of shares outstanding that have been sold short. We use the three month change instead of the one month change to capture longer-term changes should short sellers have difficulty in borrowing shares or in the event that these traders mask trades by building short positions over time.

Except for the dependent variable, regressions 4 and 5 mirror regressions 1 and 2. In regression 4, we find that the three month change in short interest is unrelated to the magnitude of the large (negative) return. In regression 5, which includes the one-day return, book-to-market ratio, natural log of shares outstanding, prior six months' return, and the standard deviation of returns, we observe a significant negative coefficient for shares outstanding (similar to regression 2). However, in this regression and unlike findings in regressions 2 and 3, the prior return is unrelated to the dependent variable. While the book-to-market ratio and standard deviation of returns do not appear to influence the *level* of short interest, the *change* in short interest is positively related to these variables. Higher book-to-market ratios (undervalued firms) and higher variations in returns (riskier

firms) appear to lead to larger increases in short interest. So, collectively, it appears that short sellers target firms with fewer shares, lower relative valuations, and greater uncertainty in returns.

Table 8: Regression Analysis

	Coefficient Estimates for Regression Number:					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	**6.939** (0.0000)	**25.274** (0.0000)	**26.982** (0.0000)	**1.274** (0.0168)	**12.120** (0.0008)	**8.721** (0.0123)
Return	4.918 (0.1285)	4.203 (0.1984)		1.318 (0.5412)	1.580 (0.4625)	
BM		2.714 (0.1974)			**2.591** (0.0625)	1.922 (0.1389)
Shares		**-1.814** (0.0002)	**-1.820** (0.0002)		**-1.187** (0.0003)	**-1.028** (0.0006)
Prior Return		**-0.715** (0.0602)	**-0.759** (0.0403)		-0.110 (0.6587)	
Std Dev		11.854 (0.5435)			**25.883** (0.0450)	19.822 (0.1019)
D0			-2.093 (0.2420)			1.757 (0.1230)
D1			-1.599 (0.3546)			1.374 (0.2124)
D2			-0.098 (0.9620)			**2.994** (0.0227)
D4			**4.7053** (0.0627)			**7.513** (0.0000)
F-statistic	2.328 (0.1285)	**3.768** (0.0027)	**4.807** (0.0001)	0.374 (.5412)	**4.638** (0.0005)	**7.740** (0.0000)
Adjusted R^2	0.006	0.059	0.094	-0.003	0.077	0.177

The dependent variable for regressions (1), (2) and (3) is the level of short interest (short sales as a percentage of shares outstanding) in the month of the large return, and the dependent variable for regressions (4), (5) and (6) is the three-month change in short interest ending in the month of the return date. Return is the "extreme" proportionate stock price change for the sample company, BM is the book-to-market ratio measured approximately six months prior to the return date, Shares represents the natural log of shares outstanding immediately prior to the return date, Prior Return is the compounded return that occurs over the six month period before the extreme return, and Std Dev is the standard deviation of the daily returns over the six month period immediately prior to the extreme return. D0, D1, D2 and D4 are dummy variables equal to one if the event leading to the extreme return is "no news," "earnings related," "strategic change," and "miscellaneous" respectively (categories are described in Table 5). p-values appear in parentheses and coefficients in bold type are statistically significant.

For regression 6, we include the variables with significant coefficients from regression 5 along with the dummy variables included in regression 3. As shown, while the shares outstanding coefficient remains negative and significant, in the presence of the event dummy variables, both the book-to-market and standard deviation coefficients are statistically insignificant. Not surprising, given the findings in Panel A of Table 7, the dummy variable coefficients for event 2 (strategic changes) and event 4 (miscellaneous events) are both positive and significant, indicating that there are larger changes in short interest for these firms prior to the negative return date. Thus, after controlling for fundamental (publicly available) information, it appears that short sellers are able to anticipate (at least partially) some negative events. Similar to Mercado-Mendez et al. (2006), these events do not appear to be directly related to earnings announcements. Based on the short interest we observe for large positive returns, however, it appears that a significant amount of short selling results from either speculative or hedging activities as opposed to informed trading.

CONCLUSIONS

Our analysis indicates that short sellers may be able to anticipate some negative news events as evidenced by the changes in short interest surrounding certain events that cause large negative returns. Specifically, we identify firms that have large one-day (negative and positive) returns and determine whether the level of or changes in short interest around this return are consistent with informed trading. Our initial evidence indicates that there are significant increases in short interest prior to negatively perceived earnings announcements and events that denote a strategic change for the firm. When we control for other factors, such as the book-to-market ratio, shares outstanding, and the standard deviation of prior returns, only strategic change and random, non-recurring events appear to be associated with significant increases in short interest prior to the large negative return.

Surprisingly, we find many cases in which the level of short interest is higher for firms experiencing large positive returns than for firms with large negative returns. Further, we find no significant declines in short interest preceding the large positive return. Thus, collectively, while we find some evidence consistent with informed trading by short sellers, the evidence is not strong. It appears that much short interest is driven by either speculation or hedging activities.

REFERENCES

Chan, W. S. (2003). Stock Price Reaction to News and No-News: Drift and Reversal after Headlines. *Journal of Financial Economics, 70*, 223-260.

Christophe, S., M. Ferri & J. Angel (2004). Short Selling Prior to Earnings Announcements. *Journal of Finance, 59*, 1845-1875.

Conrad, J. (1986). *The Price Effect of Short Sale Restrictions: Some Empirical Tests*. Ph.D. Dissertation: University of Chicago.

Desai, H., K. Ramesh, R. Thiagarajan & B. Balachandran (2002). An Investigation of the Information Role of Short Interest in the Nasdaq Market. *Journal of Finance, 57*, 2263-2287.

Diamond, D. & R. Verrecchia (1987). Constraints on Short Selling and Asset Price Adjustment to Private Information. *Journal of Financial Economics, 18*, 277-311.

Fabozzi F. & F. Modigliani (1992). *Capital Markets–Institutions and Instruments*. Englewood Cliffs: Prentice Hall.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance, 56*, 1533-1597.

Kadiyala, P. & M. Vetsuypens (2002). Are Stock Splits Credible Signals? Evidence from Short Interest Data. *Financial Management, 31*, 31-49.

Mercado-Mendez, J., R. J. Best & R. W. Best (2006). Earnings Surprise and the Relative Information Content of Short Interest. *Advances in Quantitative Analysis of Finance and Accounting, 4*(Winter), 121-135.

Reed, A. (2003). Costly Short-Selling and Stock Price Adjustment to Earnings Announcements. Working Paper: University of North Carolina.

Senchack, A. & L. Starks (1993). Short Sale Restrictions and Market Reactions to Short-Interest Announcements. *Journal of Financial and Quantitative Analysis, 28*, 177-194.

THE EFFECTS OF EXPECTATION FORMATION ON DETECTING UNEXPECTED NON-CHANGES IN ACCOUNT BALANCES DURING ANALYTICAL PROCEDURES

Charles P. Cullinan, Bryant University
Susan B. Hughes, University of Vermont

ABSTRACT

The effectiveness of audit analytical procedures in identifying financial statement misstatements has been studied from a variety of perspectives. Most analytical procedures research uses unexpected changes in an account balance or ratio to signal a possible misstatement. Some recent misstatements (e.g., Cendant, WorldCom) have involved financial statement accounts remaining the same when they should have changed (i.e., an unexpected non-change). This research assesses whether auditors forming expectations are more sensitive to unexpected non-changes in account balances than auditors not forming expectations. Seventy nine auditors from three different firms participated in an experimental analytical procedure that contained both an unexpected change and an unexpected non-change in a client's accounts. Results indicate that explicit expectation formation increases auditor's sensitivity to unexpected changes in accounts. In the unexpected non-change situation, however, there was no difference between auditors forming expectations and those not forming expectations. Those with more experience were more sensitive to the risk implications of unexpected non-changes.

Keywords: Analytical procedures, expectation formation, audit experience.

INTRODUCTION

U.S. auditing standards (AICPA, 1998) require auditors to perform analytical procedures to identify account balances that may be at a higher risk for material misstatement. As part of the analytical procedures, auditors are expected to form expectations of the balance in each key account. The effectiveness of analytical procedures has been the subject of some recent discussion, (e.g., Cullinan and Sutton, 2002; Weill, 2004) and the Securities and Exchange Commission (SEC) (1998a) has expressed concern about the appropriate use of expectation formation in analytical

procedures. The SEC specifically noted that auditors may fail to appropriately assess the risk of misstatement when balances do not change from year to year:

> ... the [SEC] staff has noted [audits in which] balances were compared from one year to the next, and as long as the balance did not change, the auditor noted that no further work was considered necessary. Often such an audit procedure is inadequate in light of ongoing changes to the company's business plans, strategies and industry conditions.

Analytical procedures have been examined from a variety of perspectives within the accounting literature. Much of this literature deals with the auditor's sensitivity and response to, changes in financial statement accounts from period to period (e.g., Mueller and Anderson 2002; Church et al. 2001). While changes in financial statement accounts can signal a potential misstatement, situations in which an account balance remains relatively static from period to period can also signal potential misstatements. This occurs when the account should change due to changes in business condition and/or recent company initiatives. Auditor insensitivity to non-changes in account balances has been noted in two recent audit failures. The former management of the Cendant Corporation manipulated its financial statements in such a way that the financial statement balances did not change enough to draw additional auditor scrutiny (SEC 2000). In the WorldCom misstatement case, the auditor, noting that "line costs as a percentage of revenue have remained flat at 41.9% of revenue . . . concluded that residual audit risk was none." (Andersen, 2002, p.12). The line cost accrual was actually misstated by approximately $3.8 billion.

Our study is motivated by the recent instances of financial statement misstatements (e.g., Cendant and WorldCom) related to non-change in financial statement accounts and/or ratios. McDaniel and Kinney (1995) concluded that an expectation formation process can enhance auditor sensitivity to potential misstatements involving an unexpected change in a financial statement account; our study investigates whether expectation formation can enhance auditors' sensitivity to potential misstatements associated with account balances that do not change when they should change (referred to in this paper as an unexpected non-change).

To identify the impact of expectation formation on unexpected non-change account balances, auditors from three CPA firms participated in an experiment in which they were asked to assess the extent of possible misstatement associated with each of the financial statement accounts included in a fictionalized case. The materials were based on the misstated financial statements filed by Aurora Foods (SEC 2001). The case materials included one account (trade and promotions allowances) that did not change even though the case materials described changes in the firm's policies in this area. The case materials also included a material decrease in interest expense that should have stayed relatively constant. To determine if forming expectations sensitized auditors to these unexpected non-changes, some respondents were asked to form expectations, while others were not.

Results indicate that, consistent with McDaniel and Kinney (1995), expectation formation increased auditors' sensitivity to unexpected ***changes*** in account balances (i.e., the unexpected decrease in interest expense). However, formal expectation formation did not increase auditors' sensitivity to the unexpected ***non-change*** in account balances (i.e., the relatively static trade and promotions allowances). In this situation, however, experience was found to be significantly associated with the auditors' sensitivity to the risk associated with the account balance.

The remainder of this paper is organized as follows. Section 2 reviews the background literature and presents the research hypotheses. Section 3 discusses the research methods and nature of the respondents. Section 4 presents the results of testing, followed by section 5 interpreting the results and providing suggestions for future research.

LITERATURE REVIEW AND HYPOTHESES

There is a rich academic literature examining auditor analytical procedure processes. Most of this literature focuses on how auditors develop and evaluate possible explanations when they have identified a misstatement risk associated with an unexpected change in account balance (e.g., Yip-ow and Tan 2000; Asare and Wright 1997). This stream of research (often relating to Ahypothesis generation@) helps to assess how auditors resolve concerns which have been detected through analytical procedures. The hypothesis generation literature expects the auditor to have identified the risk based on a material change in account balance and/or ratio. Figure 1 summarizes how some of the existing hypothesis generation literature indicated misstatement risk to their subjects. Note that in all cases, the indicator of risk is a change or fluctuation in the item of interest. McDaniel and Kinney (1995) examined the role of expectation formation guidance in analytical procedures. They find that auditors who developed formal expectations of account balances were more likely to recognize the risk that exists in an account that changed unexpectedly. They suggest that the mechanism behind their findings is task decomposition, in which the task of analytical procedures is broken down into an expectation formation phase, then a comparison of unaudited results with the expectations. They results indicate that, when requested to form explicit expectations, the auditors' focus on the expectation formation task enhanced auditor's performance in the analytical procedures task. Consistent with McDaniel and Kinney's (1995) findings, we hypothesize that:

H1: Explicit expectation formation will enhance the effectiveness of analytical procedures when there is an unexpected change in an account .

Many types of misstatements in financial statement may be signaled by a change in account balance and/or ratio such as those in indicated in Figure 1. Some recent misstatements, however, have resulted in financial statement accounts and ratios not changing when they should have. For example, in the Cendant (SEC 2000) and WorldCom (Andersen 2002, Cullinan 2004) misstatements, the financial statements were misstated because an account (expenses in these two

cases) that should have changed did not. In the audit of Perry Drug, the partner on the engagement based his audit conclusion on analytical evidence that the cost of goods sold ratio did not change, although he had physical evidence that inventory had decreased, which should have increased cost of goods sold SEC (1998b). These non-change types of misstatements may be more difficult for an auditor to detect because of the lack of a clearly recognized change signal.

Auditors should gain a firm grounding in their client's business and changes in operations for the year under audit. Based on this knowledge, the auditor can develop expectations about what should happen to the client's accounts. For a situation in which a client's account balance should change, forming more explicit expectations may cause an auditor to formally recognize that the account balance should change. Evidence from the SEC (1998a), however, indicates that auditors often form implicit expectations that account balances in financial statements should be about the same as last year. In a circumstance in which changes in a client's business *should* result in a change in an account, (say an increase in expenses), an implicit expectation of "no change" may lead auditors to miss a misstatement in which the balance did not change but should have. Formalizing the expectations process may help the auditor to overcome the implicit expectation and recognize that the account should change. In such a circumstance, the auditor would be more likely to recognize the risk of misstatement associated with an unexpected *non*-change. To assess this possibility, we test the following hypothesis:

H2: Explicit expectation formation about account changes will enhance the effectiveness of analytical procedures when there is an unexpected non-change in an account balance.

RESEARCH METHOD

Case narrative and research instrument.

The case that the respondents read and responded to was based on the Aurora Foods misstatement case (SEC, 2001; GAO, 2002). This misstatement resulted from Aurora Foods' failure to change one of the assumptions critical to the trade and promotions allowances accrual. The company misstated the "tail assumption", the amount of time between the date of the sale and the time Aurora's customers would take an allowed discount for promotional activities.

Respondents were informed they were conducting the audit planning analytical review procedures, and were presented with a case narrative and last year's audited financial statements. These financial statements were adapted from the financial statements of Aurora Foods.[1] The case narrative described a fictionalized version of Aurora Foods called "Quabbin Foods". The narrative included background information on Quabbin's operations and current year activities, and brief

biographies of the management team. The initial draft of the case was reviewed by a number of people familiar with analytical procedures, and was modified based on their input.

Figure 1: Brief Summary of Potential Misstatement Signals in Existing Analytical Procedures Literature

Paper	Potential misstatement signal
Ismail & Trotman (1995)	"change in financial ratios" (p. 349)
Yip-ow & Tan (2000)	"increase in both ... gross profit & gross profit margin" (p. 209)
Asare & Wright (1997)	"unexpected fluctuation " (p. 737)
Bierstaker, et al. (1999)	"projected* v. unaudited" (p. 22)
Anderson & Koonce (1998)	"Observed fluctuation" (p. 1)
Church, et al. (2001)	3 years data "unexpected difference for most recent year" (p. 70)
Bhattacharjee et al. (1999)	"unexpected fluctuations" (p. 87)
Mueller & Anderson (2002)	"inventory turnover ratio had significantly decreased" (p. 167)
Bedard & Biggs (1991)	projected* v. unaudited (p. 622)
Hirst & Koonce (1996)	During planning phase, auditor's most common procedure is "current year unaudited financial statements are compared to prior year audited balances" (p. 463)
Anderson, et al. (1992)	inventory turnover "exhibited a decline of five percent"
O'Donnell (2002)	"unexpected interperiod change in an account balance" (p. 179)
Anderson et al. (2003)	"unusual fluctuation" (p. 1)
* Projected based on "previous quarters of current year, past audited results, and industry trends"	

To test Hypothesis 1, we materially decreased Quabbin's interest expense. Because the company's long-term debt remained approximately the same, interest expense should have remained about the same. Thus, interest expense represents an unexpected change.

To examine hypothesis 2, the narrative indicated that as part of the company's strategic initiatives during the year under audit, Quabbin Foods (the company) "has emphasized a market push strategy to increase in-store promotions. To encourage more in-store displays of its products, Quabbin extended the time period during which its grocery store customers could take credits on their purchases to create in-store point of sale promotions from 30 to 45 days." As such, the trade and promotions on the financial statements should have increased.[2] The trade and promotions allowances for the current year were manipulated within the current year financial statements to remain relatively constant from that of the prior year, thus representing an unexpected *non*-change in financial statement account.

After reading the case narrative, all subjects were presented with Quabbin's 2002 financial statements (i.e., last year's financial statements). The first group of respondents was asked to go

directly to misstatement risk assessment. The other group of respondents was asked to form expectations, based on the narrative and last year's financials, about changes in individual financial statement accounts for the following year. These expectations were documented on a five point scale ranging from significantly decrease to significantly increase, with the mid-point being no change.

All respondents then received Quabbin's 2003 unaudited financial statements. The financial statements were presented together with the 2002 audited financials. From 2002 to 2003, the sales account changed less than one percent, while the interest expense account decreased by 40% and the trade and promotions allowances changed by less than one percent. The presentation included both dollar and common size (i.e., percentage) figures for both sets of financial statements. The presentation format of the current year unaudited financial statements compared to last year's audited financial statements is consistent with the most common procedures used during planning phase analytical procedures (Hirst and Koonce, 1996, p. 463).

Based on the comparison of the two years' financial statements, respondents were then asked to assess the "probability of misstatement based on analytical procedures". All respondents were asked to assess misstatement risk on a 5-point scale ranging from "lowest" to "highest" risk, with "moderate" risk being the mid-point. The respondent's final task was to complete a questionnaire asking about their experiences in public accounting and in conducting analytical procedures.

Respondents

Respondents were obtained from 3 regional CPA firms. Two of these firms (Firm A and Firm B) are regional CPA firms based in medium-sized cities in the Midwestern United States. The third firm (Firm C) is a regional firm based in a medium sized city in the northeastern United States that also has a practice office in a large west-coast city. For firm A, the respondents were part of a continuing professional education (CPE) session. Twenty-two responses were obtained using this method.

For firms B and C, the research instruments (case narrative, questionnaires, etc.) were distributed using the firm's interoffice mail system. Included with the research instrument was a letter from the researchers indicating that the study was interested in analytical procedures, and a supporting letter from a representative of the firm. All responses were returned directly to the researchers. Using this distribution method, 164 research instruments were distributed. Usable responses received using this distribution method were 57 for a usable response rate of 35%.

The demographic profile of the respondents is presented in Table 1. The median respondent has 5 years experience in public accounting, spends 9 of 12 months performing audit tasks, and performs 8 analytical procedures in a typical year. As such, the median respondent has performed analytical procedures 40 times (5 years times 8 analytical procedures per year) and is considered to have a reasonable level of background knowledge to complete the research instrument.

Table 1: Demographic statistics of respondents			
Continuous variables:			
Variable	Median	Mean	Std. deviation
Years experience in public accounting	5	8.7	9.0
Years with present firm	4	6.3	7.3
Months of each year performing audits	9	7.9	4.4
Number of analytical procedures performed each year	8	21.2	32.3
Discrete variables:			
Rank:	% of respondents		
Staff	45.5%		
Senior	19.0%		
Manager	26.6%		
Partner	7.6%		
Other	1.3%		
CPA Status:	% of respondents		
CPA	70.3%		
Not CPA	29.7%		

Data definitions

There are two dependent variables. The two dependent variables are the relative misstatement probabilities for the interest expense, and for trade and promotion expense. The relative misstatement probability was computed by subtracting the subjects' assessments of the misstatement probability of the sales account from the subjects' misstatement probability of the interest expense, and trade and promotions accounts, respectively. The use of relative misstatement recognizes that different respondents may perceive different base levels of risk during audit planning; the subtraction of the sales risk response measures the respondents' perception of the risk relative to the other accounts. Also the allocation of audit resources in response to risk assessment is done within clients on a relative risk basis. The sales account was chosen as the base account risk level for two reasons: (1) sales is the account most often misstated (Beasley et al. 1999), and (2) sales is the first account that the respondents were asked to rate in terms of its misstatement risk.

The independent variable of interest is the expectations variable. Expectations is a binary variable that takes the value of 1 if the respondents formed expectations or evaluated management (as appropriate) and 0 if they did not perform these tasks. Years of experience and months each year performing audits were provided by the respondents, and are included as control variables.

Testing techniques

The first step in our analyses was to compute mean relative misstatement risk for the unexpected change accounts (interest expense) and the unexpected non-change account (trade and promotions allowances). T tests were then performed comparing the misstatement risk assessment of those respondents forming expectations and those not forming expectations (providing preliminary results on hypotheses 1 and 2). Regression analyses were then computed with both of the misstatement risk assessments (for interest expense and trade and promotions allowances) as the dependent variables, and expectation, years of experience, and months of the year doing audits as the independent variables.

RESULTS

The results of testing the interest expense account are shown in Panel A of Table 2 and in Table 3. From Table 2, the mean relative misstatement risk of interest expense for those forming expectations is -0.769, while those forming expectations had a higher mean relative misstatement risk assessment of -0.098. The t value of 1.31 for this difference is significant at the 0.10 level, providing some support for hypothesis one. The regression results presented in Table 3 also indicate that respondents forming expectations had marginally significant higher relative misstatement risk assessment than those not forming expectations. Years of experience, and months devoted to audits were not significantly associated with relative misstatement risk. In summary, hypothesis 1 is supported by these interest expense results.

Table 2: Mean Relative Risk Assessments

Panel A: Mean interest expense misstatement probability (relative to sales misstatement probability)			
No expectation formation	Expectation formation	t value	p > t (one tailed)
-0.769	-0.098	-1.31	0.098
n = 39	n = 41		
Panel B: Mean trade and promotion expense misstatement probability (relative to sales misstatement probability)			
No expectation formation	Expectation formation	t value	p > t (one tailed)
-0.051	-0.125	0.22	0.914
n = 39	n = 40		

Results of testing the misstatement risk for trade and promotions allowances are shown in Panel B of Table 2 and Table 4. Table 2 shows that the mean relative misstatement risk for trade and promotions allowances for those not forming expectations was -0.051, while for those forming

expectations the mean misstatement risk was -0.125. The t value of 0.22 indicates that this difference is not statistically significant. Regression results for the trade and promotions account are presented in Table 4. Consistent with the bivariate t tests, the expectation variable is not significantly associated with relative misstatement risk for trade and promotion allowances. In contrast to the interest expense model, years of experience is significantly associated with trade and promotion risk assessments. Overall, hypothesis 2 is not supported.

Table 3: Regression Results

Dependent variable: Misstatement probability: Interest expense (relative to sales misstatement probability)

Variable:	Parameter estimate	t	p > t (one tail)*
Intercept	-0.252	-0.35	0.729*[3]
Expectation	0.663	1.33	0.094
Years experience	0.037	1.21	0.115
Months of each year doing audits	-0.106	-1.71	0.954
Model:			
F statistic	2.92		
P > F	0.039		
Adjusted R^2	0.069		
Sample size	79		

* Two tailed test due to lack of predicted direction

Table 4: Regression Results

Dependent variable: Misstatement probability: trade and promotion expense (relative to sales misstatement probability)

Variable:	Parameter estimate	t	p > t (one tail)*
Intercept	-0.067	-0.14	0.889[3]*
Expectation	-0.072	-0.22	0.914
Years experience	0.036	1.77	0.041
Months of each year doing audits	-0.036	-0.89	0.812
Model:			
F statistic	2.04		
P > F	0.116		
Adjusted R^2	0.038		
Sample size			

LIMITATIONS, DISCUSSION AND IMPLICATIONS

Limitations and suggestions for future research

This research is subject to a number of limitations. Among these limitations are potential generalizability issues, and the compromises necessary to create a research instrument sufficiently long to be realistic, and sufficiently succinct to encourage a reasonable response level. Generalizability of our results could be affected by our use of regional CPA firms, and our use of expense accounts in the experimental manipulations. Whether CPAs in international (i.e., Big 4) firms would react in a similar manner is not known. Also unknown is whether similar results would be obtained if sales and/or balance sheet accounts were subject to experimental manipulation. In a practice situation, auditors would have more sources of information than a narrative and last year's financial statements. They would have access to management and a wider array of client documents. The extent to which this broader information base may affect expectations and/or analytical review procedure evaluation is quite difficult to assess in an experimental context.

There are a number of issues relating to expectation formation and analytical procedures which are currently unexplored. Future research could examine how expectation formation may affect analytical procedures in the substantive testing and final review phases of the audit. Future research could also examine whether expectation formation causes the average perceived risk level across all accounts to increase. Also, future research could consider the effects of expectation formation on auditors' confidence in their risk assessment assessments.

Discussion of results and implications

Results of testing the interest expense misstatement risk are supportive of the results of McDaniel and Kinney (1995) and suggest that their findings that expectation formation sensitizes auditors to potential misstatements may be generalizable to regional (i.e., non-big 4) accounting firms. In contrast to the interest expense results (which indicate an unexpected change), our trade and promotion allowance results do not suggest that expectation formation improves the effectiveness of analytical procedures in all circumstances. When an account should change and does not, expectation formation did not lead auditors to be more sensitive to misstatement risk. In this situation, years of experience of the subjects lead to greater recognition of the underlying risk.

Overall, results suggest expectation formation may be important in identifying unexpected changes in financial statement accounts. When faced with unexpected non-changes, however, experience appears to be more important than explicit expectation formation. As such, the use of an expectation formation decision aid in analytical procedures may not be equally effective in alerting auditors to all types of misstatements. Implications of these results are that the effective application of analytical procedures depends on both expectation formation and performance of analytics by more experienced personnel.

ENDNOTES

[1] Because a change in accounting standards (EITF 01-9) occurred between the date of the Aurora Foods case and the date the experiments took place, the placement of trade and promotions allowances was moved from expenses (where it appeared in the original Aurora Food's financial statements) to a deduction from sales used to determine net sales (in the version of the financial statements used for this study).

[2] The mean expected change in trade and promotions allowances in the group that formed expectations (n = 41) was 3.927 on a five point scale, with 3 indicating no change. The t value for testing whether the 3.927 sample result was significantly different from 3 (no change) was 7.91, ($p < 0.0005$). The experimental manipulation appears to have been successful.

[3] * Two-tailed test due to lack of direction prediction.

REFERENCES

American Institute of Certified Public Accountants (AICPA). (1998). *Auditing Practice Release: Analytical Procedures.* New York: AICPA.

Andersen. (2002). Report to the audit committee, year ended December 31, 2001, Worldcom. Accessed through ARe: WorldCom, Inc., HO-09440 Revised Statement Pursuant to Section 21(a)(1) of the Securities Exchange Act of 1934" at 8/23/04.

Anderson, J. C., S. E. Kaplan & P. M. J. Reckers. (Fall 1992). The effects of output interference on analytical procedure judgments. *Auditing: A Journal of Practice and Theory, 11*(2), 1- .

Anderson, J. C., K. K. Moreno & J. M. Mueller. (2003). The effect of client vs. decision aid as a source of explanations upon auditor's sufficiency judgments: A research note. *Behavioral Research in Accounting, 15*, 1-11.

Anderson, U. & L. Koonce. (1998). Evaluating the sufficiency of causes in audit analytical procedures. *Auditing: A Journal of Practice and Theory 17*(1), 1-12.

Asare, S. K. & A. Wright. (1997). Hypothesis revision strategies in conducting analytical procedures. *Accounting, Organizations and Society, 22*(8), 737-755.

Beasley, M., J. Carcello & D. Hermanon. (1999). *Fraudulent Financial Reporting: 1987-1997 An Analysis of U.S. Public Companies.* (New York: Committee of the Sponsoring Organizations of the Treadway Commission).

Bedard, J. C. & S. F. Biggs. (1991). Pattern recognition, hypothesis generation, and auditor performance in an analytical task. *Auditing: A Journal of Practice and Theory , 66*(3), 622-642.

Bhattacharjee, S., T. Kida & D. M. Hanno. (Spring, 1999). The impact of hypothesis set size on the time efficiency and accuracy of analytical review judgments. *Journal of Accounting Research, 37*(1), 83-100.

Bierstaker, J. L., J. C. Bedard & S. F. Biggs. (Spring, 1999). The role of problem representation shifts in auditor decision processes in analytical procedures. *Auditing: A Journal of Practice and Theory, 18*(1), 18-36.

Church, B., J. McMillan & A. Schneider. (March, 2001). Factors affecting internal auditors' consideration of fraudulent financial reporting during analytical procedures. *Auditing: A Journal of Practice and Theory, 20*(1), 65-80.

Cullinan, C. (2004). Enron as a symptom of audit process breakdown: Can the Sarbanes-Oxley Act cure the disease. *Critical Perspectives on Accounting, 15*, 853-864.

Cullinan, C. & S. Sutton. (June 2002). Defrauding the public interest: A critical examination of reengineered audit processes and the likelihood of fraud detection. *Critical Perspectives on Accounting, 13*, 297-310.

General Accounting Office (GAO) (October, 2002). Financial Statement Restatements: Trends, Market Impacts, Regulatory Responses and Remaining Challenges. Washington: GAO.

Gullapalli, D. (2004). In the grocery world, rebates remain the mystery meat. *The Wall Street Journal*, 8/9/04, C1.

Hirst, D. E. & Koonce, L. (Fall, 1996). Audit analytical procedures: A field investigation. *Contemporary Accounting Research 13*(2), 457-486.

Ismail, Z. & K. T. Trotman. (1995). The impact of the review process in hypothesis generation tasks. *Accounting, Organizations and Society, 20*(5), 345-357.

McDaniel, L. S. & W. R. Kinney, Jr., (1995). Expectation formation guidance in the auditor's review of interim financial information. *Journal of Accounting Research,* 33(1), 59-76.

O'Donnell, E. (2002). Evidence of association between error-specific experience and auditor performance during analytical procedures. *Behavioral Research in Accounting, 14*, 179-195.

Securities Exchange Commission (SEC). (1998a). Commission Information: Letter From the Office of the Chief Accountant Regarding the 1998B1999 Audit Risk Alerts. Accessed at 8/23/04.

Securities Exchange Commission (SEC). (1998b). In the matter of Richard Valade, CPA, respondent. Accessed at http://www.sec.gov/litigation/admin3440002.txt (accessed August 23, 2004).

Securities Exchange Commission (SEC). (2000). In the matter of Cendant Corporation, respondent (AAER 1272). Accessed at 8/23/04.

Securities Exchange Commission (SEC). (2001). In the matter of Linda Mueller, Respondent (AAER 1362). Accessed at: 8/23/04.

Weill, J. (2004). Behind a wave of corporate fraud: A change in how auditors work. *Wall Street Journal*. March 25, A1.

Yip-ow, J. & H-T. Tan. (2000). Effects of preparer's justification on the reviewer's hypothesis generation and judgment in analytical procedures. *Accounting, Organizations and Society, 25*, 203-215.

EARNINGS MANIPULATION TO ACHIEVE COGNITIVE REFERENCE POINTS IN INCOME

Charles E. Jordan, University of Southern Mississippi
Stanley J. Clark, University of Southern Mississippi
Gwen R. Pate, University of Southern Mississippi

ABSTRACT

Thomas (1989) demonstrates that U.S. firms with positive earnings manipulate income by rounding up the second earnings digit to increase the first earnings digit by one, thereby reaching cognitive reference points in income. Companies with negative income manipulate earnings in the opposite direction. The current study replicates Thomas' (1989) twenty-year old research to ascertain if this type of earnings management continues today, particularly in light of the heightened scrutiny managers now face to present fair financial reporting. The results suggest that managers of negative earnings firms no longer engage in this manipulative behavior while managers of positive earnings firms continue to do so. In addition, for positive earnings firms, the findings indicate that the propensity to engage in this form of earnings management is related to specific company characteristics. Small firms appear to manage earnings to achieve cognitive reference points more intensively than large firms. Likewise, low-leverage companies exhibit this manipulative behavior more frequently than high-leverage firms, and less profitable firms engage in this activity more aggressively than companies enjoying high profit margins.

INTRODUCTION

Although no unique definition exists for earnings management, its primary attribute embodies the manipulation of the financial reporting process to create personal gain (Jackson & Pitman, 2001). Managers manipulate earnings for many reasons, and specific lines of research have developed around each form of earnings management. For example, earnings are manipulated to smooth income and, thereby, create the impression that a firm is low-risk to investors or creditors (Ronen & Sadan, 1981) or to meet analysts' earnings expectations, thus maintaining a steady growth in share price (Payne & Robb, 2000; Jordan & Clark, 2003). Earnings are typically manipulated upward but may be managed downward as firms take big hits to income in years when earnings are already depressed (Peek, 2004; Yoon & Miller, 2002). In addition, managers manipulate income to avoid reporting losses or decreases in earnings (Burgstahler & Dichev, 1997) or to enhance executive bonuses tied to income (Healy, 1985; Guidry et al., 1999).

A particular subset of earnings management research initiated by Carslaw (1988) examines the notion that firms manage earnings upward to reach cognitive reference points in income. Brenner and Brenner (1982) indicate that humans possess only a limited amount of memory and that memory is used to store the most relevant pieces of information about a price. As an example, in the price $496 more emphasis will be placed on the first digit (4), less on the second digit (9), and so forth. In remembering numbers, people tend to round down rather than up. For example, $496 would be rounded down to $490 or even to $400 rather than rounded up to the more logical $500.

Carslaw (1988) speculates that managers manipulate income to just above key cognitive reference points if the unmanaged earnings fall only slightly below the reference points. As an example, earnings of $891 million would be managed upward to an amount slightly above $900 million. As a result, financial statement users in their rounding would round down to $900 million rather than to $800 million. Carslaw (1988) and others demonstrate that this type of manipulative behavior exists as unusual digital patterns are observed in the second-from-the-left earnings position for firms with positive earnings (i.e., in particular, more zeros and fewer nines than expected appear in the second earnings position).

The current article extends this line of research by replicating it with recent data on U.S. companies to ascertain if earnings are still manipulated to achieve cognitive reference points in income despite the increased scrutiny managers now face to present fair financial reporting. In addition, specific company characteristics that have been linked to the propensity to manage earnings (i.e., firm size, Glaum et.al., 2004; debt leverage, Sweeney, 1994; and operating performance, Bauman et al., 2001) are investigated to determine if they are associated with the likelihood of firms manipulating earnings to achieve cognitive reference points in income. Results show that this form of earnings management continues for U.S. firms with positive income and that the propensity to engage in this manipulative behavior relates to all three firm traits examined (i.e., entity size, leverage, and operating performance).

LITERATURE REVIEW

The current study examines the earnings management literature from two perspectives. First, articles analyzing earnings manipulation to achieve cognitive reference points in income are reviewed. Second, studies examining the association between company characteristics and the propensity to manage earnings are discussed.

As noted previously, Carslaw (1988) is the first researcher to examine the notion that management manipulates earnings to reach cognitive reference points in income. His premise is that managers will find ways to increase the left most position (i.e., first digit) in earnings by one if unmanaged earnings are already close to the breakpoint. If such earnings manipulation occurs, Carslaw (1988) surmises that the digital patterns of the second earnings position would not follow expected frequencies. He hypothesizes that this second earnings position would be made up of an

abnormally high frequency of low digits (e.g., zeros and ones) and an unusually low proportion of high digits (e.g., eights and nines).

Crucial to this type of research is specifying an appropriate expected frequency of digits in each earnings position. Intuition suggests that each of the digits one through nine possesses a one-ninth chance of occurring in the first position of an earnings number. Then, each of the digits zero through nine would have a one-tenth chance of appearing in earnings positions to the right of the first digit. However, this is not the case. Carslaw (1988) relies on work by Feller (1966) to obtain his expected frequencies. Feller's (1966) expected digital frequencies are identical to those developed by Benford (1938), which is now commonly referred to as Benford's law.

Benford (1938) notes that, for naturally occurring numbers, the first position in a number possesses a large bias for low digits. For example, ones appear in the first position about 31% of the time with twos occurring in this position approximately 18% of the time. Using integral calculus, Benford (1938) computes the expected frequencies of the digits in various positions within a number. He tests his expected frequencies against 20 widely diverging lists of numbers with a total of 20,229 observations. Included in his list of numbers are items such as the geographic areas of rivers, a list of all the numbers in an issue of *Reader's Digest*, and the first 342 street addresses given in the *American Men of Science*. Benford finds that his expected digital frequencies accurately reflect the observed digital frequencies for numbers that are without known relationships. Benford's law does not apply to numbers that follow an orderly fashion, such as postal zip codes or invoice numbers, nor does it work with small numbers (e.g., ages of individuals in a group). Nigrini (1999) shows that Benford's expected digital frequencies are well suited for identifying manipulative biases in financial statement numbers. Table 1 provides Benford's expected digital frequencies for the first two positions within a number.

Table 1: Benford's Expected Digital Frequencies

	Position of digit in number	
Digit	First	Second
0		11.97%
1	30.10%	11.39
2	17.61	10.88
3	12.49	10.43
4	9.69	10.03
5	7.92	9.67
6	6.70	9.34
7	5.80	9.04
8	5.12	8.76
9	4.58	8.50

Source: Nigrini (1999, p. 80).

Using the expected frequencies shown in Table 1, Carslaw (1988) examines the second position in the earnings numbers for publicly traded New Zealand companies with positive income. He finds that there exists a much higher than expected frequency of zeros and lower than expected proportion of nines in the second earnings position. All other digits (i.e., one through eight) in this second position conform to Benford's expected frequencies. Carslaw (1988) concludes that this anomaly occurs because managers adjust the income numbers upward to increase the first earnings digit by one when the second digit is close to the breakpoint (i.e., income is adjusted to increase the second digit from nine to zero, thereby increasing the first digit by one).

Thomas (1989) replicates Carslaw's (1988) study with data for publicly traded U.S. firms. His primary goal is to discover if the digital patterns noted in the earnings figures for New Zealand companies by Carslaw (1988) hold true for U.S. firms. In addition, Thomas (1989) examines the income numbers for both positive and negative earnings firms, while Carslaw (1988) investigates only positive earnings firms. Thomas (1989) examines the second earnings position for all firms on the 1986 edition of the Annual Industrial and Over the Counter COMPUSTAT files. Like Carslaw (1988), he compares the actual frequencies of the digits appearing in the second earnings position to Benford's expected frequencies (see Table 1). Thomas (1989) employs the normally distributed Z-statistic to determine whether the deviations between the observed and expected frequencies of digits are statistically significant. Similar to the New Zealand firms in Carslaw's (1988) research, Thomas (1989) finds that U.S. companies with positive earnings report more zeros and fewer nines than expected in the second earnings position. He also shows that negative earnings firms exhibit just the opposite pattern (i.e., fewer zeros and more nines than anticipated). Thus, Thomas (1989) concludes that firms with profits manipulate the second earnings position to reach a one-number-higher digit in the first position while companies with losses manage the second digit to achieve a one-number-lower digit in the first position.

Niskanen and Keloharju (2000) replicate the Carslaw (1988) and Thomas (1989) studies with Finnish firms. Examining only positive earnings firms, their results generally confirm those of the earlier researchers but provide evidence that the manipulation of the second earnings digit to achieve a higher first digit is even greater for the Finnish companies. In particular, Niskanen and Keloharju (2000) find that the four highest digits (i.e., six, seven, eight, and nine) each appear in the second earnings position less often than expected. Interestingly, though, only sixes and sevens exhibit observed frequencies that are different from their expected proportions at statistically significant levels. Their results also show significantly higher than anticipated frequencies of both zeros and ones in the second position. With these findings, Niskanen and Keloharju (2000) surmise that the earnings management practiced by Finnish firms is far more than merely cosmetic. Finnish firms do not limit their manipulation or rounding up of the second digit merely from nines to zeros but do so from sixes and sevens all the way to zeros and ones.

Using U.K firms with positive earnings, Van Caneghem (2002) replicates the Carslaw (1988), Thomas (1989), and Niskanen and Keloharju (2000) studies. He examines pre-tax earnings and reaches conclusions similar to the previous studies. In particular, Van Caneghem (2002) finds

that when managers are faced with a nine in the second earnings position they manipulate earnings upward to round the second earnings digit to zero, thereby increasing the first digit by one. The primary contribution of the Van Caneghem (2002) study to this line of research is its demonstration that management uses short-range, discretionary accruals to accomplish the upward movement in earnings. Van Caneghem (2004) examines audit quality to determine if it represents a factor in constraining managers from rounding the second earnings digit to achieve a one-number-higher first digit. Again using a sample of U.K. firms, he finds no difference in the propensity of big-five versus non-big-five auditors to inhibit this form of earnings manipulation by managers.

The second line of earnings management research relevant for the current project involves studies examining the relation between company characteristics and the propensity to manage earnings. For example, Glaum et al. (2004) and Aharony (1993) conclude that company size is associated with the propensity to manage earnings, with smaller firms manipulating income more intensely than larger ones. On the other hand, Jordan et al. (2007) provide evidence that larger firms are more likely to manage earnings than smaller ones. Glaum et al. (2004) note that the relation between firm size and the propensity to manage earnings is unclear because of the different reasons that firms manipulate income.

The level of debt leverage has also been linked to the propensity to manage earnings. Companies with higher debt leverages represent more risky investments. Thus, to compensate the owners for the risks they are taking, Watts and Zimmerman (1986) speculate that managers of these highly leveraged companies may be more prone to manipulate earnings upward. In addition, heavily leveraged firms that are close to violating debt covenants or are on the brink of financial distress may be more likely to manage earnings than are companies with low debt levels (Aharony et al., 1993; Sweeney, 1994; Porcano, 1997; Visvanathan, 1998).

The likelihood of managers manipulating earnings has been tied to operating performance. Yoon and Miller (2002) demonstrate that firms manage income upward if earnings performance is poor. Similarly, Burgstahler and Dichev (1997) suggest that companies manipulate earnings upward to prevent reporting negative profits or declining earnings. Yet, companies with negative earnings or earnings substantially depressed relative to prior periods may take large discretionary provisions in the current period to "clear the decks" for future periods (Bauman et al., 2001).

The prior literature reviewed sets the stage for the current research. Only one study involving the manipulation or rounding of earnings to achieve cognitive reference points in income examines U.S. firms (i.e., Thomas, 1989), and it does so with 1986 data. In recent times, management misbehavior has come under heightened scrutiny in the U.S. because of notable business failures linked to fraudulent financial reporting (e.g., Enron, WorldCom, Tyco, etc.). The SEC, Congress, courts, investing public, and external auditors have all become less tolerant of unethical management behavior, and the Sarbanes-Oxley Act now provides stiff penalties for fraudulent financial reporting. Thomas' (1989) study is replicated in the current research using 2006 data on U.S. firms to investigate whether managers continue to manipulate earnings to achieve cognitive reference points in income in spite of the increased pressure to present unbiased financial reporting.

The intersection of the research on earnings management to reach cognitive reference points in income with the literature on the association between company characteristics and the propensity to manipulate earnings leads to the second issue addressed in the present study. In particular, for the U.S. companies sampled, the rounding of earnings to achieve reference points in income and specific company characteristics (i.e., firm size, leverage, and operating performance) are jointly examined to determine if relationships exist between these characteristics and the likelihood that companies will engage in this form of earnings management.

METHODOLOGY

To address the research issues above, selected 2006 financial statement data are collected for 1002 randomly selected, publicly traded U.S. companies. The data are obtained from the LexisNexis Business database. Because one purpose of the current research is to replicate Thomas' (1989) study using recent data, information is obtained for both positive and negative earnings firms. The concern here is whether U.S. firms continue to exhibit the same earnings manipulation behavior that Thomas (1989) observes in his 1986 data. More specifically, do positive earnings firms manipulate income by rounding the second earnings digit up to increase the first earnings digit by one while negative earnings firms round the second digit down to decrease the first digit by one?

Consistent with prior research (i.e., Carslaw, 1988; Thomas, 1989; Niskanen & Keloharju, 2000; Van Caneghem, 2002), the earnings position of interest in the current study is the second earnings digit. If positive earnings firms round earnings to increase the first digit by one, the second position should possess a lower than expected frequency of high digits (e.g., fewer eights and nines) and a higher than expected proportion of low digits (e.g., more zeros and ones). The opposite result would be anticipated if this type of manipulation occurs with negative earnings firms.

To keep the earnings figures as free from noise as possible, only companies without discontinued operations and extraordinary items are examined. The income figure analyzed represents bottom-line earnings. The prior studies in this area all assume that, in the absence of earnings management, the observed frequencies of digits in the second earnings position should conform to Benford's (1938) expected proportions. If the observed frequencies of digits in the second position do not conform to Benford's expected frequencies, strong anecdotal evidence exists that earnings are managed (i.e., rounded) to achieve cognitive reference points in the first earnings position. The current study applies this same methodology.

The second issue addressed is whether the propensity to manage earnings to reach cognitive reference points in income is related to certain company characteristics (i.e., firm size, debt leverage, and operating performance). The analysis here concentrates on the second earnings position for positive earnings firms only. No analysis is made of negative earnings firms because these companies have less incentive to achieve specific levels of income. As Burgstahler and Dichev (1997) note, a primary reason firms manage income is to avoid earnings decreases or losses. Once a firm's earnings go negative or it fails to meet its earnings expectations, the market punishes it

relatively the same in terms of reduced share price, regardless of whether it misses its earnings mark by a little or a lot (Jordan & Clark, 2003). Thus, if a negative earnings firm engages in earnings manipulation it is more likely to be manipulation geared toward taking a big hit to income to "clear the decks" for future earnings. Little regard would be paid to achieving a particular earnings reference point in the current year (i.e., a year with earnings already depressed). In their examination of earnings rounding to reach cognitive reference points in income, Van Caneghem (2002), Niskanen and Keloharju (2000), and Carslaw (1988) analyze strictly positive earnings firms; the current study does likewise in investigating the relation between earnings manipulation and firm size, leverage, and operating performance.

To determine if the propensity to round earnings to achieve a one-number-higher first digit in earnings is related to firm traits (i.e., size, leverage, or operating performance), the group of positive earnings firms is segregated into two subgroups based on the particular characteristic examined. Using the debt ratio (i.e., total debt to assets) to measure leverage, the group of positive earnings firms is split into two subgroups with one containing high-leverage firms and the other made up of low-leverage companies. For each subgroup, the observed frequencies of digits in the second earnings position are compared to the expected (i.e., Benford's) frequencies. If the actual digital patterns differ from the expected frequencies for one subgroup but not the other subgroup or if both subgroups' digital patterns differ from the expected frequencies but do so in unique fashions, then evidence exists indicating that the propensity to round earnings to achieve cognitive reference points in income is related to debt leverage. Similar analyses are made by dividing the group of positive earnings firms into subgroups based both on firm size and operating performance.

RESULTS

As noted previously, using 1986 data, Thomas' (1989) study of U.S. companies shows that positive earnings firms report more zeros and fewer nines than expected in the second earnings position while negative earnings firms present fewer zeros and more nines than anticipated in this position. Thus, he concludes that positive earnings firms round income up to achieve a one-number-higher first earnings digit but negative earnings firms round the earnings figure down to present a first digit that is one number lower. Table 2 provides an analysis of the observed and expected frequencies for the second earnings position for the 749 positive earnings firms and 253 negative earnings firms in the current study (i.e., using 2006 data for U.S. companies).

Using a .05 cutoff level for statistical significance, Table 2 reveals no unusual digital patterns for the second earnings position for negative earnings firms. The observed frequencies for all digits (i.e., zero through nine) conform to Benford's expected proportions. Thus, unlike Thomas' (1989) results from two decades earlier, negative earnings firms no longer appear to manipulate income to reach cognitive reference points in the first earnings position. However, the same cannot be said for positive earnings firms.

For positive earnings firms, the observed frequency of every high digit (i.e., five, six, seven, eight, and nine) in the second position is lower than expected, although none of the deviations individually is statistically significant at a .05 level. The two lowest digits (i.e., zero and one) occur much more frequently than expected, and with significance levels of .050 and .009, respectively, these deviations are not due to random chance. The results suggest that positive earnings firms manipulate income to round up the second earnings digit to increase the first digit by one.

The fact that negative earnings firms no longer seem to manipulate income to reach cognitive reference points in income does not necessarily mean that this activity ended due to the heightened pressure managers now face to present fair and transparent financial reporting, but this represents one plausible explanation. However, one fact is clear; even with today's increased diligence of the SEC, courts, Congress, and external auditors to produce a financial reporting system free from manipulative biases, managers of positive earnings firms continue to intervene in the reporting process to bring about desired outcomes that are favorable to their companies.

Table 3 presents the digital frequencies for the second earnings position for the group of 749 positive earnings firms segregated into two subgroups by firm size. Firm size is measured by total assets; the median total assets for the entire group of positive earnings firms is $736 million. The 375 companies with total assets below this median comprise the subgroup of small firms, while the firms with total assets above $736 million make up the subgroup of large firms.

For the small firms, a clear digital pattern occurs that suggests earnings manipulation. In particular, there are fewer high digits (i.e., sevens, eights, and nines) than expected in the second earnings position, with the deviation for nines displaying a significance level of .044. Conversely, there are more low digits (i.e., zeros and ones) than expected, with the deviation for ones statistically significant at a .001 level. For the large firms, though, the digital patterns are not so definitive. Although there are fewer nines and more zeros and ones than expected in the second position, the deviations are not statistically significant (i.e., at a .05 cutoff level). Thus, the information in Table 3 suggests that the propensity to manipulate earnings to effect cognitive reference points in income is related to firm size, with small firms exhibiting a much stronger tendency to engage in this behavior than large ones.

Table 4 provides the observed and expected digital frequencies for the second earnings position for the group of positive earnings firms separated into two subgroups by debt leverage. The median debt ratio (i.e., debt to assets) for the entire group of 749 positive earnings firms is 52.09%. The 375 firms with debt ratios below this median comprise the subgroup of low-leverage firms, while the 374 companies with debt ratios above 52.09% make up the high-leverage subgroup.

The subgroup of low-leverage firms exhibits clear signs that earnings are managed to increase the first earnings digit by rounding up the second digit. In particular, fewer sevens, eights, and nines than expected appear in the second earnings position, with the deviation for eights possessing a significance level of .031. In addition, more zeros and ones than anticipated occur in the second position, even though the deviation for neither digit is statistically significant at a .05 level. Thus, the low-leverage firms demonstrate a classic pattern of earnings rounding to achieve

cognitive reference points in income (i.e., fewer high digits and more low digits than expected in the second earnings position, with at least one of the deviations statistically significant).

Table 2: Digital Frequencies for Negative and Positive Earnings Firms

	Second Earnings Digit									
	0	1	2	3	4	5	6	7	8	9
Negative Earnings Firms (n=253):										
Observed count (n)	27	21	26	32	29	24	26	24	24	20
Observed freq. (%)	10.67	8.30	10.28	12.65	11.46	9.49	10.28	9.49	9.49	7.91
Expected freq. (%)	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
Z value	-.635	-1.546	-.309	1.153	.758	-.098	.514	.250	.410	-.339
P level	.525	.122	.757	.249	.449	.922	.607	.802	.681	.738
Positive Earnings Firms (n=749):										
Observed count (n)	107	108	70	83	70	65	68	65	63	50
Observed freq. (%)	14.29	14.42	9.35	11.08	9.35	8.68	9.08	8.68	8.04	6.68
Expected freq. (%)	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
Z value	1.954	2.611	-1.350	.581	-.624	-.917	-.243	-.341	-.335	-1.790
P level	.050	.009	.177	.562	.532	.359	.808	.733	.738	.073

Note: The expected frequency for each digit is based on Benford's (1938) law. Z value and p level relate to a one-sample proportions test comparing the observed frequency of a digit to its expected frequency.

The high-leverage subgroup also shows signs of earnings rounding to reach cognitive reference points in income, but to a lesser degree than the low-leverage subgroup. The high-leverage firms possess fewer nines and more zeros and ones than anticipated in the second earnings position, yet none of the deviations is statistically significant at a .05 level. A primary difference in the results between the subgroups segregated by the level of debt is that the low-leverage subgroup apparently rounds the second earnings digit over a wider range than the high-leverage subgroup. In addition, none of the deviations between the observed and expected frequencies of digits in the second earnings position is statistically significant for the high-leverage firms, while a high digit (i.e., eight) appears much less often than should occur by chance for the low-leverage subgroup. Thus, the level of debt seems to be related to the propensity that firms will manipulate earnings to achieve cognitive reference points in income, with low-leverage firms engaging in this behavior more aggressively than high-leverage firms.

Table 3: Digital Frequencies for Small and Large Firms

	Second Earnings Digit									
	0	1	2	3	4	5	6	7	8	9
Small Firms (n=375):										
Observed count (n)	52	64	34	36	34	35	40	29	30	21
Observed frequency (%)	13.87	17.07	9.07	9.60	9.07	9.33	10.67	7.73	8.00	5.60
Expected frequency (%)	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
Z value	1.133	3.461	-1.129	-.528	-.622	-.219	.885	-.879	-.519	-2.014
P level	.257	.001	.259	.598	.534	.826	.376	.379	.604	.044
Large Firms (n=374):										
Observed count (n)	55	44	36	47	36	30	28	36	33	29
Observed frequency (%)	14.71	11.77	9.63	12.57	9.63	8.02	7.49	9.63	8.82	7.75
Expected frequency (%)	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
Z value	1.631	.229	-.780	1.350	-.261	-1.08	-1.230	.398	.046	-.517
P level	.103	.819	.435	.177	.794	.281	.219	.690	.964	.605

Note: The expected frequency for each digit is based on Benford's (1938) law. Z value and p level relate to a one-sample proportions test comparing the observed frequency of a digit to its expected frequency.

Table 5 presents the digital frequencies for the second earnings position for the group of positive earnings firms segregated by operating performance. Return on assets (ROA), computed as net income to total assets, measures operating performance; the median ROA for the 749 positive earnings firms is 5.97%. The 375 companies with ROAs below 5.97% make up the subgroup of low-ROA firms, while the 374 firms with ROAs exceeding 5.97% comprise the high-ROA subgroup.

An intriguing finding occurs with the low-ROA subgroup. At first glance, it appears that no earnings manipulation exists for this subgroup as the observed frequency for the second earnings position for no digit differs from its expected proportion at a statistically significant level (i.e., .05). However, every high digit (i.e., five, six, seven, eight, and nine) appears less often than expected, while every low digit (i.e., zero, one, two, three, and four) occurs more frequently than anticipated.

Such a digital pattern likely results from liberal earnings manipulation similar to that observed by Niskanen and Keloharju (2000) for the Finnish firms. That is, earnings manipulation is not confined simply to rounding up very high second digits (like eight or nine) just across the breakpoint to zero or one. Instead, the rounding extends for the entire length of the digital range for the second digit.

Table 4: Digital Frequencies for Low and High Leverage Firms

	Second Earnings Digit									
	0	1	2	3	4	5	6	7	8	9
Low Leverage Firms (n=375):										
Observed count (n)	53	54	33	47	37	30	43	33	21	24
Observed frequency (%)	14.13	14.40	8.80	12.53	9.87	8.00	11.47	8.80	5.60	6.40
Expected frequency (%)	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
Z value	1.292	1.835	-1.295	1.331	-.106	-1.093	1.417	-.162	-2.163	-1.458
P level	.196	.066	.195	.183	.916	.274	.156	.871	.031	.145
High Leverage Firms (n=374):										
Observed count (n)	54	54	37	36	33	35	25	32	42	26
Observed frequency (%)	14.43	14.43	9.89	9.63	8.82	9.36	6.68	8.56	11.23	6.95
Expected frequency (%)	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
Z value	1.472	1.856	-.614	-.511	-.777	-.203	-1.763	-.323	1.692	-1.074
P level	.141	.064	.539	.610	.437	.839	.078	.747	.091	.283

Note: The expected frequency for each digit is based on Benford's (1938) law. Z value and p level relate to a one-sample proportions test comparing the observed frequency of a digit to its expected frequency.

Table 5: Digital Frequencies for Low and High ROA Firms

	Second Earnings Digit									
	0	1	2	3	4	5	6	7	8	9
Low ROA Firms (n=375):										
Observed count (n)	46	50	41	49	38	35	30	31	32	23
Observed frequency (%)	12.27	13.33	10.93	13.07	10.13	9.33	8.00	8.27	8.53	6.13
Expected frequency (%)	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
Z value	.178	1.185	.032	1.668	.066	-.219	-.890	-.519	-.153	-1.643
P level	.859	.236	.975	.095	.947	.826	.374	.604	.878	.100
High ROA Firms (n=374):										
Observed count (n)	61	58	29	34	32	30	38	34	31	27
Observed frequency (%)	16.31	15.51	7.75	9.09	8.56	8.02	10.61	9.09	8.29	7.22
Expected frequency (%)	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
Z value	2.587	2.508	-1.943	-.849	-.949	-1.078	.547	.038	-.320	-.888
P level	.010	.012	.052	.396	.342	.281	.584	.970	.749	.374

Note: The expected frequency for each digit is based on Benford's (1938) law. Z value and p level relate to a one-sample proportions test comparing the observed frequency of a digit to its expected frequency.

With rounding apparently not limited to a few specific digits for the low-ROA subgroup, it seems these firms should be evaluated using frequencies for groups of digits rather than the individual digits' frequencies. More specifically, the high digits (i.e., five, six, seven, eight, and nine) represent one group of digits with the other group comprising the low digits (i.e., zero, one, two, three, and four). The combined observed frequency of high digits is 40.27%, and the combined expected frequency (i.e., summing Benford's expected proportions) of the high digits is 45.30%. The z value and p-level for the deviation in these two proportions is –1.957 and .050, respectively. For the low digits, the combined observed frequency is 59.73% while the combined expected frequency is 54.70%. The z value and p-level for the difference between these two frequencies is

1.957 and .050, respectively. Simply stated, the low-ROA firms possess too few high digits and too many low digits in the second earnings position than should occur by chance. This result highly suggests that the low-ROA subgroup liberally rounds the second earnings digit to achieve a cognitive reference point in the first digit.

Table 5 shows that the high-ROA subgroup also exhibits signs of earnings manipulation as there exist fewer eights and nines in the second position than expected, even though the deviations between the actual and expected proportions are not statistically significant at a .05 level. In addition, this subgroup possesses more zeros and ones than expected in the second position, and these differences are statistically significant (i.e., significance levels of .010 and .012, respectively, for zeros and ones). Appreciably more zeros and ones than expected in the second position implies that high-ROA firms round earnings up to just above the breakpoint in the second position to enable the first digit to increase by one. Thus, both the low-ROA and high-ROA subgroups appear to manage earnings to achieve cognitive reference points in income. However, they accomplish their manipulations in different manners with high-ROA firms apparently limiting the rounding of the second position to a few specific digits surrounding the breakpoint while low-ROA firms seem much more aggressive in their manipulations (i.e., rounding over a much wider range of digits).

SUMMARY AND CONCLUSION

The present study addresses two research issues. First, considering the current heightened scrutiny of managers and pressures on them to present unbiased financial reporting, does earnings manipulation to achieve cognitive reference points in income continue as it did twenty years ago at the time of Thomas' (1989) research? Our results indicate that negative earnings firms no longer engage in this type of earnings management. Whether managers of negative earnings firms discontinued this activity because of the increased scrutiny is unknown, but this certainly represents one reasonable explanation for their change in behavior. For positive earnings firms, though, the evidence suggests that managers continue to round up the second earnings digit to increase the first digit by one. These results do not necessarily mean that managers of positive earnings firms have not changed their behavior in response to the increased scrutiny; however, the findings do suggest that this heightened scrutiny has failed to eliminate earnings manipulation to achieve cognitive reference points in income.

The second research issue examined is whether the propensity to manipulate earnings to achieve cognitive reference points in income appears to be associated with company characteristics (i.e., firm size, leverage, and operating performance). The results suggest that all three firm traits appear to be related to the likelihood that firms will manage earnings in this manner.

Large firms exhibit no strong tendencies to engage in this manipulative behavior, while small firms seem to embrace it. Perhaps as Glaum et al. (2004) note, small firms may be more likely than large ones to manage earnings because it is more difficult for them to raise capital. Thus, small firms may manage earnings more intensely to increase their attractiveness in the capital markets. Another

plausible explanation for the discrepancy between large and small firms is that, as Watts and Zimmerman (1986) suggest, managers of large firms face more public scrutiny than managers of small firms. Perhaps the Sarbanes-Oxley Act and news of corporate executives at major corporations receiving jail time for fraudulent financial reporting produce more of a deterrent for misbehavior for the managers of large firms than they do for the executives of small ones. Of course, this is merely speculation as the current study's purpose is only to examine whether a relation exists between firm size and the propensity to manipulate earnings to reach cognitive reference points in income. Future research would be needed to ascertain why the association occurs.

Although both high and low leverage firms exhibit some signs of rounding earnings to achieve cognitive reference points in income, only the low-leverage subgroup possesses digital patterns suggesting that this manipulative behavior is excessive. This finding contradicts prior research (e.g., Aharony et al., 1993; Visvanathan, 1998) indicating that high-leverage firms are more prone to earnings management than are low-leverage companies. However, the earlier studies examining the relation between leverage and earnings management did not investigate the rounding of earnings to reach cognitive reference points in income. Perhaps the low-leverage firms cannot obtain higher levels of debt financing because they are viewed as less credit worthy for some reason and are, therefore, rounding earnings up to achieve breakpoints in income to appear more credit worthy. It is also possible that some factor other than debt leverage is driving the results. For example, firms in the low-leverage subgroup are generally much smaller than companies in the high-leverage subgroup. The median total assets for the low-leverage firms of $399 million is less than one-third of the median total assets for the high-leverage firms of $1.27 billion, and these medians differ at a statistically significant level (i.e., p value = .000). As noted previously, small firms appear aggressive in their rounding of the second earnings digit to achieve a cognitive reference point in the first digit while large companies show no strong signs of engaging in this manipulative behavior. Perhaps the size factor overrides leverage and is causing the low-leverage subgroup to appear more prone to earnings manipulation than the high-leverage subgroup. Again, though, these theories are untested and represent only a couple of plausible explanations for this outcome. Future research would be needed to determine why low-leverage firms seem to round earnings to achieve cognitive reference points in income more intensively than companies with high debt levels.

With ROA representing the measure of operating performance, both high and low performing firms seem to manipulate earnings to reach cognitive reference points in income, but they do so in unique fashions. More profitable firms appear somewhat conservative in their rounding of earnings and confine it to situations where the second earnings digit is relatively high (i.e., eight or nine) and then manipulate earnings just enough to cross the breakpoint to increase the first digit by one. Less profitable companies, however, seem to engage in much more liberal manipulations by rounding up any high second digit (i.e., five, six, seven, eight, or nine), and they do not limit the rounding to just across the breakpoint. This finding is intuitive as logic suggests that high performing firms have less need to manipulate earnings than low performing companies. As Yoon and Miller (2002) note, when operating performance is poor companies select strategies to

improve earnings. Thus, it is not surprising to see that the less profitable companies in this study manage earnings more aggressively than the more profitable ones.

REFERENCES

Aharony, J., C. Lin, & M.P. Loeb (1993). Intial public offerings, accounting choices, and earnings management. *Contemporary Accounting Research*, 10(1), 61-81.

Bauman, C.C., M.P. Bauman, & R.F. Halsey (2001). Do firms use the deferred tax asset valuation allowance to manage earnings? *Journal of American Taxation Association*, 23(supplement), 27-48.

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, 78(4), 551-572.

Brenner, G.A. & R. Brenner (1982). Memory and markets, or why are you paying $2.99 for a widget. *The Journal of Business*, 55(1), 147-158.

Burgstahler, D. & I. Dichev (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics*, 24(1), 99-126.

Carslaw, C. (1988). Anomalies in income numbers: Evidence of goal oriented behavior. *The Accounting Review*, 63(2), 321-327.

Degeorge, F., J. Patel, & R. Zeckhauser (1999). Earnings management to exceed thresholds. *Journal of Business*, 72(1), 1-33.

Feller, W. (1966). *An introduction to probability theory*. New York: Wiley & Sons.

Glaum, M., K. Lichtblau, & J. Lindeman (2004). The extent of earnings management in the U.S. and Germany. *Journal of International Accounting Research*, 3(2), 45-77.

Guidry, F., A.J. Leone, & S. Rock (1999). Earnings-based plans and earnings management by business-unit managers. *Journal of Accounting and Economics*, 26(1-3), 113-142.

Healy, P.M. (1985). The effect of bonus schemes on the selection of accounting principles. *Journal of Accounting and Economics*, 7(1-3), 85-107.

Jackson, S. & M. Pitman (2001). Auditors and earnings management. *The CPA Journal*, 71(7), 39-44.

Jordan, C.E. & S.J. Clark (2003). Evidence on the level of corporate America's participation in the earnings game. *The National Accounting Journal*, 5(1), 61-70.

Jordan, C.E, S.J. Clark, & M.A. Waldron (2007). Company characteristics and the use of SFAS No. 143 to effect earnings management. *Academy of Accounting and Financial Studies Journal*, 11(2), 97-108.

Nigrini, M.J. (1999). I've got your number. *Journal of Accountancy*, 187(5), 79-83.

Niskanen, J. & M. Keloharju (2000). Earnings cosmetics in a tax-driven accounting environment. *European Accounting Review*, 9(3), 443-452.

Payne, J.L. & S.W.G. Robb (2000). Earnings management: The effect of ex ante earnings expectations. *Journal of Accounting, Auditing and Finance*, 15(4), 371-392.

Peek, E. (2004). The use of discretionary provisions in earnings management: Evidence from the Netherlands. *Journal of International Accounting Research*, 3(2), 27-43.

Porcano, T.M. (1997). An analysis of capital gains tax-induced earnings management. *International Advances in Economic Research*, 3(4), 395-408.

Ronen, J. & S. Sadan (1981). *Smoothing income numbers: Objectives, means and implications*. Wokingham: Addison-Wesley

Sweeney, A.P. (1994). Debt-covenant violations and managers' accounting responses. *Journal of Accounting and Economics*, 17(3), 281-308.

Thomas, J.K. (1989). Unusual patterns in reported earnings. *The Accounting Review*, 64(4), 773-787.

Van Caneghem, T. (2002). Earnings management induced by cognitive reference points. *British Accounting Review*, 34(2), 167-178.

Van Caneghem, T. (2004). The impact of audit quality on earnings rounding behavior: Some U.K. Evidence. *European Accounting Review*, 13(4), 771-786.

Visvanathan, G. (1998). Deferred tax valuation allowances and earnings management. *Journal of Financial Statement Analysis*, 3(4), 6-15.

Watts, R.L. & J.L. Zimmerman (1986). Towards a positive theory of the determination of accounting standards. *The Accounting Review*, 53(1), 112-134.

Yoon, S.S. & G.A. Miller (2002). Cash from operations and earnings management in Korea. *International Journal of Accounting*, 37(4), 395-412.

THE EARNINGS IMPACT OF FAS 154: AN ANALYSIS BY YEAR, INDUSTRY AND FIRM SIZE

Douglas K. Schneider, East Carolina University
James M. Kohlmeyer, East Carolina University
Mark W. Rieman, East Carolina University

ABSTRACT

When a new accounting standard is proposed a question arises as to whether or not the impact on earnings will be significant. The recently issued FAS 154 mandates that enterprises no longer report the cumulative-effect losses and gains of a change in accounting method. Based on 1998 through 2004 data, the removal of the cumulative-effect loss or gain from the income statement would significantly impact reported earnings.

INTRODUCTION

In May 2005 the Financial Accounting Standards Board (FASB) issued Statement of Financial Accounting Standards No. 154, *Accounting Changes and Error Corrections*—a replacement of *APB Opinion No. 20* and *FASB Statement No. 3*, (FAS 154) (FASB 2005a). The new standard is part of a joint effort on the part of the FASB and the International Accounting Standards Committee to create greater convergence between U.S. and international accounting standards (FASB 2005b). FAS 154 is a substantially different method of reporting the effects of a change in accounting principle and correction of errors than its predecessor rule, Accounting Principles Board Opinion No. 20 (APB Opinion 20; AICPA 1971).

APB Opinion 20 required that in the year an entity changed accounting methods the effects of that change be reported on the income statement as a cumulative-effect gain or loss on prior years of that change in accounting principle' (here in "cumulative-effect"). The cumulative-effect gain or loss was reported as a special net-of-tax amount at the bottom of the income statement, after income from continuing operations. In contrast FAS 154 does not require nor allow the cumulative-effect to be reported on the income statement. Instead, the effects of changes in accounting principle are accounted for by "retrospective application." Under retrospective application there is no special reporting feature to the income statement in the year of change. Instead, the current-year income statement will reflect the new method. On the annual report however, any previous years' financial statements presented will show the "retrospective application" of the new method as if it had been in effect during those years. Thus, the previous years presented will be more comparable to the current year in which the new accounting method was first used.

Since FAS 154 eliminates the cumulative-effect gain or loss, we suggest that FAS 154 will have an impact on reported net income. The cumulative-effect on prior years' gain or loss required under APB Opinion 20 was the overall earnings effect on all prior years, not just the effect on the current year's net income for the year of the accounting change. Therefore, in the year an enterprise changed accounting methods the earnings impact on the income statement was potentially significant. Our research question asks, "What is the likely impact under FAS 154 of no longer reporting the cumulative-effect gain or loss?" This question is explored by examining the impact on net income if FAS 154 had been applied to fiscal years 1998-2004.

Itemizing or cataloging the specific changes in accounting principle for the over 2,600 firm-observations used in this study is not the focus of this study and is beyond its scope. However, some specific new accounting standards issued by the FASB appear to account for a substantial portion of the mandatory accounting changes that occurred during the years under study. The Accounting Trends & Techniques (AICPA 2003) revealed that during 1999-2002 fourteen different new accounting standards or Emerging Issue Task Force rulings became mandatory. This increase in cumulative-effect accounting changes appeared to mostly reduce earnings.

For an example, in 2001 the FASB issued Statement of Financial Accounting Standards No. 143, *Accounting for Asset Retirement Obligations* (FAS 143) (FASB 2001). This accounting standard requires that an enterprise recognize and measure the cost of the obligation to retire an asset at the end of its useful life as well as recognize imputed interest expense and the incremental amount of depreciation expense. One hundred forty-one of the 600 firms surveyed in Accounting Trends & Techniques reported this as an accounting change. Other frequent mandatory changes during this time period involved accounting for business combinations, goodwill, stock compensation, debt extinguishments, guarantees, derivatives and consolidation of variable interest.

In addition to required changes in accounting principle, discretionary changes appeared to involve changes in accounting for inventory and changes in revenue recognition.
Also, during the down years in the stock market after 2000 some firms may have engaged in isolated discretionary accounting changes that had the affect of a "big bath" write-off of earnings.

This study finds evidence that if FAS 154 had been in effect from 1998 through 2004, it would have resulted in higher earnings for most enterprises. In addition, this study conducted t-tests to compare the distribution of the percentage impact on earnings of FAS 154 between the firms that had reported cumulative-effect gains and those that had reported cumulative-effect losses. From the results of this study it is hoped that inferences can be drawn as to the anticipated impact of FAS 154 on future years.

PREVIOUS RESEARCH

Accounting changes has been the subject of numerous accounting studies. Some earlier studies examined accounting changes from the perspective of the apparent motive of discretionary changes and if firms would be more likely to make voluntary accounting changes if earnings were

increased than if earnings were decreased. Cushing examined the impact of APB Opinion No. 20 and found noncompliance with the accounting rules (Cushing 1974). May and Schneider (1988) found some evidence suggesting that discretionary accounting changes were motivated by a desire to manage earnings.

Earlier studies exploring accounting changes often focused on the stock market's response to an accounting change announcement. Harrison (1977) reported that both discretionary and nondiscretionary accounting changes that increase net income are associated with concurrent and unique stock market behavior. However, the negative return differences for discretionary changes, contrasted with positive return differences for nondiscretionary changes, suggest that the discretion available to management in making the accounting changes possesses information content.

Ricks and Hughes (1985) questioned the Harrison (1977) study by reexamining the market's reaction to a change from the cost to the equity method of accounting for long-term investments. They argued for two major weaknesses of Harrison (1977) study. First, there may be other events which cause adjustments of the market's expectations regarding the impact of an accounting change. For example, the market may react at times when an impending change is proposed, revised, approved, or enacted (e.g. Lev (1979), Pfeiffer (1980), Smith (1981), Hughes & Ricks (1984)). Second, there is a decrease in statistical power which results from using lengthy time periods to test for market reactions (Brown & Warner, 1980). Ricks & Hughes (1985) resolved these issues by identifying the dates when specific public disclosures of events related to the change are made and then examined price behavior near to those dates (Binder, 1983; Schipper & Thompson, 1983). The results indicated evidence to support the view that earnings adjustments precipitated by the change contained new information. However, no market reaction was detected in the weeks containing public announcements leading up to and including the APB's adoption of the change.

Abdel-khalik and McKeown (1978) incorporated earnings expectations into an explanatory model of the equity security market's response to announcements of a discretionary accounting policy change. Cready, Shank, and McKeown (1987) replicated Abdel-khalik and McKeown's (1978) study and attempted to improve the experimental design. Results indicated that Abdel-khalik and McKeown's (1978) finding that the market's reaction to the switch to LIFO is conditional on the sign of the earnings forecast error is still supportable, although the magnitude of its effect is diminished relatively.

Chung, Park, and Ro (1996) examined whether the direction of voluntary accounting method changes for inventory, depreciation, and investment tax credit have a stock price implication. Direction of accounting changes (accounting changes) in each area is defined relative to the most commonly-used accounting practices by industry. Results show that the abnormal stock returns of the sample firms around the accounting change announcements are, on average, positive for the accounting changes away from the common accounting practices in the selected areas and negative for the accounting changes towards the practices. Evidence suggests that the direction of voluntary accounting changes relative to the common accounting practices has a stock price implication, and

that the uniformity of accounting rules across all firms may limit investors' access to some firm-specific information.

Other prior research on voluntary accounting changes indicated that voluntary accounting changes can affect stock price if they affect future cash flows via, for example, taxes (e.g. Morse and Richardson, 1983; Biddle and Lindahl, 1982; Johnson and Dhaliwal 1988) or contracting/monitoring (Holthausen and Leftwich 1983). Stock price might also be affected if voluntary accounting changes affect the precision of info (Verrecchia 1982; Holthausen and Verrecchia 1990).

Lastly, Moore, Atkinson, and Nix (2003) make an argument for eliminating the cumulative effect adjustment because it could be misleading to users. Some cumulative effects are not in the income statement but rather are carried directly to retained earnings as an adjustment to the beginning balance. Several exceptions are mentioned (e.g., change from the LIFO method on inventory pricing, change in the method of accounting for long-term construction contracts, change to or from the full cost method in the extractive industries, change to the equity method of accounting for investments).

This study adds to previous research by addressing a different and timely issue, namely the earnings impact of no longer reporting the cumulative-effect gain or loss as required by a new accounting standard.

RESEARCH ISSUES

The issuance of FAS 154 raises an obvious question. What is the potential impact on earnings of FAS 154? Enterprises and investors would have a greater concern over a new standard if it has a potentially significant impact on earnings.

Since accounting changes occurring in the future cannot be predicted with a high level of certainty, an alternative way to assess the likely potential impact of FAS 154 is to examine the earnings impact in previous years as if FAS 154 had been applied during those years. This study examined the fiscal years 1998 through 2004. For each year the enterprises were segmented into two subsets; enterprises that reported cumulative-effect losses (here-in "loss firms") and enterprises that reported cumulative-effect gains (here-in "gain firms"). The segmentation of each year's sample into loss firms and gain firms provides a greater opportunity for insights into the cumulative-effect on earnings by comparing the loss firms and gain firms.

This research design provides an overall look at the impact of FAS 154 as opposed to merely examining selected specific enterprises as anecdotal examples. Thus, a broader question is whether the adoption of the new accounting standard would, on the whole, have a pervasive impact on the population of enterprises that had reported cumulative-effect gains or losses. In other words, while the new standard could have a material affect on a few enterprises, would there be a significant overall affect for the enterprises as a group that reported cumulative-effect gains or losses.

The importance of an accounting standard can be measured in a variety of ways under different research designs. This study examines a population of enterprises that had reported

cumulative-effect gains and losses. For each enterprise the percentage impact of FAS 154 on earnings is measured as if it had been in effect for fiscal years 1998 through 2004. Descriptive statistics are generated and an analysis comparing loss firms and gain firms is conducted using both parametric and nonparametric tests.

SAMPLE AND DATA COLLECTION

The data collection for this study consists of enterprises which reported on the Compustat (or *Research Insight*) database a cumulative-effect gain or loss (data item A183) for the fiscal years 1998 through 2004. A total of 3,005 firms were initially found which reported a cumulative-effect gain or loss. However, of the initial 3,005 observations 2,612 were ultimately included in the study because 393 of the initial firms reported missing data. Three hundred-sixty firms did not report market value or shares outstanding because such firms were subsidiaries of publicly traded firms included in the study. Including the subsidiaries and their parent companies in the sample would have resulted in duplication. Thus, subsidiary firms were eliminated from the study. The remaining 33 firms were eliminated because of data missing for other variables.

The 2,612 firms included in the study are summarized by year below. The firms are segmented into those firms which reported cumulative-effect losses and those that reported cumulative-effect gains. The firms are segmented in order to provide more meaningful results.

Fiscal Year	Total Number of Cumulative-effect Firms	Number of Loss Firms	Number of Gain Firms
1998	167	141	26
1999	252	229	23
2000	409	373	36
2001	487	343	144
2002	687	637	50
2003	511	369	142
2004	99	71	28
	2612	2,163	449

The number of firms in the study varied from a high of 687 in 2002 to a low of 99 in 2004. Table 1 shows the number and types of firms that reported cumulative effects from changes in accounting principles (heretofore changes) for the years 1998 through 2004, segmented by loss firms and gain firms. There were 2,163 loss firms that reported changes from the 10 different industry classifications during this period (Panel A). The largest number of firms are in the manufacturing industry (SCI codes 20-39), 831, with 2000 through 2003 showing the greatest number. The

services industry (SIC code 70-89) shows the second largest number of loss firms, 363. Three of the industries show observations of less than 13 firms.

Table 1: Frequency Distribution by Industry (Two-digit SIC Code) and Year of Sample Firms Reporting Cumulative-effect of a Change in Accounting Principle.
Total number of firms in the sample is 2,612.

Panel A – Firms Reporting Cumulative-effect Losses:									
		Number of Firms by Year							
SIC Code	**Industry**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**Total**
01-09	Agriculture, Forestry, Fishing	1	0	1	0	4	0	0	6
10-14	Mining	3	10	11	15	15	65	2	121
15-17	Construction	0	0	0	0	8	3	0	11
20-39	Manufacturing	45	58	187	132	268	128	13	831
40-49	Transportation & Public Utilities	19	19	41	39	84	69	14	285
50-51	Wholesale Trade	5	8	4	10	30	14	2	73
52-59	Retail Trade	26	40	25	15	39	27	3	175
60-67	Finance, Insurance, Real Estate	12	47	28	86	53	28	32	286
70-89	Services	30	47	75	43	132	31	5	363
91-99	Public Administration	0	0	1	3	4	4	0	12
	Total	141	229	373	343	637	369	71	2163
Panel B – Firms Reporting Cumulative-effect Gains:									
01-09	Agriculture, Forestry, Fishing	1	2	0	0	0	0	0	3
10-14	Mining	4	0	1	13	3	51	4	76
15-17	Construction	0	3	2	2	0	0	0	7
20-39	Manufacturing	10	9	11	40	14	30	5	119
40-49	Transportation & Public Utilities	8	2	14	37	12	44	10	127
50-51	Wholesale Trade	0	2	1	5	1	0	0	9
52-59	Retail Trade	0	3	2	2	3	1	1	12
60-67	Finance, Insurance, Real Estate	1	2	1	26	13	10	6	59
70-89	Services	2	0	4	18	3	3	2	32
91-99	Public Administration	0	0	0	1	1	3	0	5
	Total	0	23	36	144	50	142	28	449

Panel B shows that among gain firms manufacturing (SIC code 20-39) and transportation & public utilities (SIC code 40-49) are, 119 and 127, respectively. For the gain firms four different

industries showed less than ten firms. Thus, accounting changes and cumulative-effect gains, and losses, were present among some industries more than others.

Table 2: Frequency Distribution by Sales Quartiles for Years of Sample Firms Reporting Cumulative-effect of a Change in Accounting Principle.

Panel A – Sales Quartiles for the Entire Sample Consolidated for 1998-2004:
The sales quartiles are determined based on the entire sample consolidated from 1998-2004. These quartiles are used to create subsets for analyzing both loss firms and gain firms.

Quartiles	Sales Range (in millions)	Loss Firms	Gain Firms	Total
1st	$0.008 - 121.890	559	94	653
2nd	$121.890 - 608.498	559	94	653
3rd	$608.498 – 2,912.209	545	108	653
4th	$2,912.209 – 213,199.000	500	153	653
Total		2,163	449	2,612

Panel B – Loss Firms by Sales Quartiles by Year:

Quartiles	Sales Range (in millions)	1998	1999	2000	2001	2002	2003	2004	Total
1st	$0.008 - 121.890	40	60	160	74	136	78	11	559
2nd	$121.890 - 608.498	45	69	104	82	173	78	8	559
3rd	$608.498 – 2,912.209	33	57	59	89	192	96	19	545
4th	$2,912.209 – 213,199.000	23	43	50	98	136	117	33	500
Total		141	229	373	343	637	369	71	2163

Panel C – Gain Firms by Sales Quartiles by Year:

Quartiles	Sales Range (in millions)	1998	1999	2000	2001	2002	2003	2004	Total
1st	$0.008 - 121.890	10	5	11	22	14	32	0	94
2nd	$121.890 - 608.498	4	6	9	33	9	25	8	94
3rd	$608.498 – 2,912.209	7	6	9	32	15	31	8	108
4th	$2,912.209 – 213,199.000	5	6	7	57	12	54	12	153
Total		26	23	36	144	50	142	28	449

Table 2 shows the distribution of firms according to quartiles by the reported sales of the sample firms. In panel B the gain firms are more weighted towards larger firms with 108 firms in the 3rd quartile and 153 firms in the 4th quartile. The loss firms are weighted more towards smaller firms with 559 firms in both the 1st and 2nd quartiles as compared to 545 and 500 firms in the 3rd and 4th quartiles.

Table 3: Reported Sales of Sample Firms
(All columns in millions of dollars except year and n.)

This table is intended to provide some indication of the size of the sample firms using reported sales as a measure of firm size.

Panel A - Loss Firms:

Year	n	Mean	Median	Standard Deviation	p25	p75	Minimum	Maximum
1998	141	3,061.468	331.753	11,453.836	94.500	1,542.372	0.185	100,697.000
1999	229	3,266.270	410.817	12,916.360	115.477	1,952.271	0.122	165,013.000
2000	373	2,467.800	194.435	10,093.561	33.648	818.446	0.057	152,446.000
2001	343	5,654.422	833.181	14,396.528	140.713	3,953.132	0.084	125,679.000
2002	637	3,855.122	635.218	12,824.161	163.994	2,362.106	0.504	162,586.000
2003	369	6,208.531	1,029.353	18,464.982	199.352	4,779.000	0.008	171,869.688
2004	71	8,984.848	1,859.308	17,222.522	430.113	7,156.400	2.488	97,152.000
Total	2,163							

Panel B - Gain Firms:

Year	n	Mean	Median	Standard Deviation	p25	p75	Minimum	Maximum
1998	26	2,674.570	440.555	6,260.613	40.148	1,706.500	8.365	28,777.000
1999	23	6,364.288	788.160	12,978.664	159.266	4,647.000	9.831	49,489.000
2000	36	1,845.533	419.887	3,650.187	77.760	1,770.702	1.661	18,781.396
2001	144	7,111.598	1,429.998	12,839.755	304.120	8,068.500	0.404	61,257.000
2002	50	3,761.760	826.390	7,863.781	95.001	2,657.872	2.672	37,878.109
2003	142	9,182.044	1,189.705	24,783.454	173.125	6,624.322	0.098	213,199.000
2004	28	5,926.561	1,804.621	8,981.353	521.003	9,021.769	218.406	42,089.000
Total	449							

To further describe the profile of the sample firms Table 3 shows the descriptive statistics of the sales of the sample of loss and gain firms by year. This provides an indication of the size of the firms that reported cumulative-effect losses and gains. For the loss firms the largest sales amount was in 2003, $171,869,688,000. The minimum amount, also in 2003, was $80,000. The median sales of the loss firms increased during the time period under study with 2003 and 2004 exceeding $1 billion. For the gain firms the median sales fluctuated with 2001, 2003 and 2004 exceeding $1 billion.

The next section covers the results of the study.

RESULTS

This section discusses the statistics conducted on the sample firms to measure the impact of FAS 154 on firms' earnings. Presented in Tables 4, 5 and 6 are the percentage change in earnings by year, industry and sales quartiles. Later, Tables 7, 8 and 9 present paired t-tests of pre- and post-FAS 154 earnings by year, industry and sales quartiles.

To measure the potential impact of FAS 154 the actual earnings of the sample firms are recalculated as if FAS 154 had been in effect for the year that a firm reported a cumulative-effect gain or loss on prior years' earnings due to a change in accounting principle. To explore the effect of FAS 154 analyses are conducted of pre-FAS 154 earnings as actually reported and post-FAS 154 earnings as they would have been reported if FAS 154 had been in effect at the time. To provide a common measure the percentage change in earnings from retroactive application of FAS 154 is used. To reveal the effect of FAS 154 the firms are segmented into those that reported cumulative-effect losses and those that reported cumulative-effect gains. Presenting results without segmenting the firms would not be particularly helpful since losses and gains have the opposite effect and tend to cancel each other in such a way that mean and medians would not be meaningful and would report understated amounts closer to zero than if the two groups of firms are examined separately.

Table 4 presents descriptive statistics by year of the percentage change in earnings if FAS 154 were retroactively applied to 1998 through 2004. In Panel A, the loss firms, the means range from 23.1% increase in 2001 to 141.2% in 2002. The median is a more conservative measure of the percentage increase in earnings if FAS 154 were retroactively applied. It ranged from a low of 3.5% in 2004 to a high of 59.5% in 2002. The 75th percentile varies from 12.9% to 108.3%. So, for 2002 one-fourth of the firms would have experienced an increase in earnings of greater than or equal to 108.3% if FAS 154 had been retroactively applied in 2002 and thus firms would not have reported cumulative-effect losses.

In Panel B the gain firms, which are fewer in number than the loss firms, show mean decreases in earnings that range from -11.7% to -43.6% and the median ranges from -2.3% to -12.4%. For the gain firms the 25th percentile shows the percentage decrease in earnings that one-fourth of the firms would have experienced in each year. The percentage change in earnings for both loss and gain firms appears to be substantial for each year.

The impact by industry is presented in Table 5. In Panel A the loss firms show a wide range in mean and median percentage change in earnings. The median percentage change in earnings is greatest for Wholesale Trade and Construction, though the number of firms were fewer than Manufacturing and Services. While the earnings impact is minimal below the 25th percentile, however the impact appears to be large for most industries at the 75th percentile. The impact does not seem quite as large for the gain firms in Panel B, however the 25th percentile shows deteriorations in earnings from a -4.1% decrease to a -40.4% decrease. Four of the industries in Panel B contain less than ten firms and Retail Trade has only 11 firms. The industries with a limited number of observations do not provide meaningful descriptive statistics and prevent useful

comparisons to other industries with more firms. However, firms in most industries did show double-digit percentage change in earnings.

Table 4: Descriptive Statistics by Year of Percentage Change in Earnings if FAS 154 had been in Effect from 1998 through 2004 (Amounts in percentages and not absolute values)

(Cumulative-effect losses and gains would have been absent from earnings if FAS 154 had been in effect during these years.)

Panel A - Loss Firms:

Year	N	Mean	Median	Standard Deviation	p25	p75	Minimum	Maximum
1998	141	39.8	9.8	111.0	3.7	33.9	0.3	1,024.8
1999	229	49.9	6.3	207.4	2.6	22.5	0.2	2,450.0
2000	373	62.4	8.0	241.8	2.4	27.2	0.1	2,500.0
2001	343	23.1	3.5	63.7	1.2	14.8	0.0	670.7
2002	637	141.2	59.5	555.9	22.6	108.3	0.3	11,633.3
2003	369	51.9	6.3	215.3	1.6	23.6	0.1	2,220.2
2004	71	50.5	3.5	245.1	1.3	12.9	0.1	2,033.4
	2,163							

Panel B - Gain Firms:

Year	N	Mean	Median	Standard Deviation	p25	p75	Minimum	Maximum
1998	26	-43.6	-12.3	81.4	-34.1	-4.0	-351.0	-0.4
1999	23	-33.2	-10.1	79.0	-25.0	-2.5	-373.3	-0.4
2000	36	-18.1	-12.4	23.1	-22.1	-2.7	-116.1	-0.1
2001	144	-11.8	-2.3	27.3	-9.1	-0.8	-197.3	0.0
2002	50	-40.9	-6.6	127.3	-28.7	-2.6	-868.7	-0.2
2003	142	-13.7	-4.8	21.2	-17.3	-1.5	-162.2	0.0
2004	28	-11.7	-4.0	16.8	-15.3	-1.3	-72.6	-0.4
	449							

Table 5: Descriptive Statistics by SIC Code of Percentage Change in Earnings if FAS 154 had been in Effect from 1998 through 2004
(Statistical amounts are in percentages and not absolute values.)

Panel A - Loss Firms:

SIC Code	Industry	N	Mean	Median	Standard Deviation	p25	p75	Minimum	Maximum
01-09	Agriculture, Forestry, Fishing*	6	81.1	22.4	128.2	3.7	104.6	3.1	330.5
10-14	Mining	121	27.8	7.1	70.2	2.1	18.9	0.1	485.5
15-17	Construction	11	47.8	29.5	40.5	6.6	95.3	1.0	103.6
20-39	Manufacturing	831	72.1	16.1	197.0	3.6	72.2	0.1	2,500.0
40-49	Transportation & Public Utilities	285	99.9	8.5	717.2	2.6	44.0	0.0	11,633.3
50-51	Wholesale Trade	73	111.7	36.9	266.9	4.4	91.1	0.3	1,571.8
52-59	Retail Trade	175	118.8	10.3	472.5	2.7	38.9	0.1	4,595.3
60-67	Finance, Insurance, Real Estate	286	39.6	4.3	115.6	1.5	19.7	0.1	1,014.1
70-89	Services	363	74.3	20.7	281.0	4.6	67.7	0.1	3,992.3
91-99	Public Administration	12	76.4	13.3	124.2	5.6	103.4	1.3	393.5
		2,163							

Panel B - Gain Firms:

SIC Code	Industry	N	Mean	Median	Standard Deviation	p25	p75	Minimum	Maximum
01-09	Agriculture, Forestry, Fishing*	1	-4.1	-4.1		-4.1	-4.1	-4.1	-4.1
10-14	Mining	78	-20.1	-4.1	39.3	-20.7	-1.3	-204.8	-0.1
15-17	Construction*	4	-21.6	-13.3	26.6	-40.4	-2.8	-59.1	-0.8
20-39	Manufacturing	113	-21.9	-6.1	85.6	-15.1	-1.7	-868.7	0.0
40-49	Transportation & Public Utilities	134	-18.1	-4.8	38.5	-18.2	-1.2	-351.0	-0.1
50-51	Wholesale Trade*	9	-53.5	-7.0	120.9	-32.7	-2.1	-373.3	-1.1
52-59	Retail Trade	11	-23.5	-15.7	35.6	-27.9	-4.1	-126.6	-0.4
60-67	Finance, Insurance, Real Estate	60	-15.1	-3.9	31.4	-14.1	-1.0	-195.2	-0.1
70-89	Services	34	-9.7	-4.4	14.0	-14.7	-1.2	-69.7	-0.2
91-99	Public Administration*	5	-3.5	-2.6	2.5	-5.5	-1.5	-6.9	-1.3
		449							

*The limited number of observations in some industries may not meet data requirements to provide meaningful results

To analyze if the size of the firm might affect the percentage change in earnings if FAS 154 were retroactively applied the firms are analyzed by sales quartiles. For the loss firms in Panel A the median percentage increase in earnings ranged from 5.3% for the 4th quartile to 17.5% for the 2nd quartile. The 75th percentile ranges from 29.6% to 63.4%. One conclusion that can be drawn is that the largest firms reported a smaller percentage increase than the three smaller quartiles.

Panel B shows the gain firms. The 1st quartile shows the largest percentage decrease in earnings, a median of -7.9% and -26.5% at the 25th percentile. The 4th quartile, the largest firms, shows in Panel B a median of -2.8% decrease in earnings and -11.0% at the 25th percentile. The largest firms show a smaller decrease in earnings than the smaller three quartiles, a pattern that exists for both loss and gain firms. So, apparently the largest firms would have experienced the smallest earnings impact and the smaller firms would have experienced a larger earnings impact.

Table 6: Analysis by Sales Quartiles Descriptive Statistics by Sales Quartiles of Percentage Change in Earnings if FAS 154 had been in effect from 1998 through 2004

(All statistical amounts are percentages and not absolute values.)

Panel A - Loss Firms:

	Sales Quartile	Sales Range (in millions)	n	Mean	Median	Std. Dev.	p25	p75	Minimum	Maximum
	1st	$0.008 - 121.890	559	63.3	16.8	170.2	4.8	63.4	0.1	2,500.0
	2nd	$121.890 - 608.498	559	82.8	17.5	278.4	3.5	62.2	0.1	3,992.3
	3rd	$608.498 – 2,912.209	545	93.3	10.3	565.7	2.5	62.0	0.0	11,633.3
	4th	$2,912.209 – 213,199.000	500	56.7	5.3	213.7	1.8	29.6	0.1	2,450.0
Total			2,163							

Panel B - Gain Firms:

	Sales Quartile	Sales Range (in millions)	n	Mean	Median	Std. Dev.	p25	p75	Minimum	Maximum
	1st	$0.008 - 121.890	94	-29.4	-7.9	55.9	-26.5	-1.7	-351.0	-0.1
	2nd	$121.890 - 608.498	94	-23.9	-5.1	92.5	-16.2	-1.7	-868.7	-0.1
	3rd	$608.498 – 2,912.209	108	-13.9	-5.3	20.1	-17.3	-1.7	-122.8	0.0
	4th	$2,912.209 – 213,199.000	153	-13.5	-2.8	36.3	-11.0	-0.9	-373.3	0.0
			449							

An additional analysis was conducted to statistically measure the extent of the earnings impact of FAS 154 on the sample firms in this study. That FAS 154, retroactively applied, would impact earnings was a fore drawn conclusion. And, it would not be surprising that the impact might be large at least for some individual firms. A larger question is the overall impact of FAS 154 on the firms as a group. To explore this issue paired t-tests are conducted that compare pre-FAS 154 (i.e., earnings as actually reported) and post-FAS 154 earnings (i.e., earnings that would have been

reported if FAS 154 had been retroactively applied). This analysis, a paired t-test, will determine if the difference in the two earnings numbers has a mean greater than zero. A practical interpretation of these results is whether FAS-154 would have significant, pervasive overall impact on earnings across the entire sample of firms. To add insight into this analysis paired t-tests results are presented by year, industry and sales quartiles.

Table 7: Paired T-test by Year Comparing Pre FAS-154 Earnings versus Post FAS-154 Earnings

Panel A - Loss Firms:				
	Year	**t Value**	**DF**	**Pr > |t|**
	1998	-1.4653	140	0.1451
	1999	-5.4410	228	0.0000
	2000	-4.4825	372	0.0000
	2001	-5.2100	342	0.0000
	2002	-2.8875	636	0.0040
	2003	-4.3583	368	0.0000
	2004	-1.9829	70	0.0513
Panel B - Gain Firms:				
	Year	**t Value**	**DF**	**Pr > |t|**
	1998	1.9836	25	0.0584
	1999	1.4373	22	0.1647
	2000	1.3909	35	0.1730
	2001	3.3738	143	0.0010
	2002	2.0336	49	0.0474
	2003	3.1151	141	0.0022
	2004	4.0622	27	0.0004

Table 7, Panel A, indicates that in five of the seven years the earnings impact of FAS 154 would have been significant. It was not significant in 1998 and was at best marginal in 2004. Panel B, the gain firms, shows that from 2001 through 2004 the difference is significant, but it was not for 1999 and 2000 and was marginal in 1998. However, 1998 through 2000 had fewer firms and lower degrees of freedom than 2001 through 2004.

The t-test results by industry are shown in Table 8 show that for loss firms in Panel A eight out of the ten industries showed significant results. One of the industries that is not significant, Agriculture, Forestry, Fishing has only five degrees of freedom due to limited observations, which may prevent meaningful conclusions. For gain firms in Panel B six of the ten industries showed

significant results. However, for both significant and insignificant industries the number of observations was limited. For both sets of firms the limited number of observations in some industries may not meet the necessary data requirements to provide meaningful results. Yet, for most of the industries in which the number of observations was not a concern the results were significant.

Table 8: T-test by Industry (SIC) Comparing Pre-FAS 154 Earnings versus Post-FAS 154 Earnings

Panel A - Loss Firms:					
	SIC	**Industry**	**t Value**	**DF**	**Pr > |t|**
	01-09	Agriculture, Forestry, Fishing*	-2.1042	5	0.0893
	10-14	Mining	-2.0134	120	0.0463
	15-17	Construction*	-2.9317	10	0.0150
	20-39	Manufacturing	-7.4204	830	0.0000
	40-49	Transportation & Public Utilities	-3.1987	284	0.0015
	50-51	Wholesale Trade	-3.1655	72	0.0023
	52-59	Retail Trade	-3.6137	174	0.0004
	60-67	Finance, Insurance, Real Estate	-3.1665	285	0.0017
	70-89	Services	-1.3246	362	0.1862
	91-99	Public Administration*	-3.7829	11	0.0030
Panel B - Gain Firms:					
	SIC	**Industry**	**t Value**	**DF**	**Pr > |t|**
	01-09	Agriculture, Forestry, Fishing*	---	---	---
	10-14	Mining	2.4150	77	0.0181
	15-17	Construction*	7.3189	3	0.0053
	20-39	Manufacturing	3.1377	112	0.0022
	40-49	Transportation & Public Utilities	3.3476	133	0.0011
	50-51	Wholesale Trade*	4.4350	8	0.0022
	52-59	Retail Trade*	2.1569	10	0.0564
	60-67	Finance, Insurance, Real Estate	2.3582	59	0.0217
	70-89	Services	1.2743	33	0.2115
	91-99	Public Administration*	1.5413	4	0.1981

*The limited number of observations in some industries may not meet data requirements to provide meaningful results

Table 9: Paired T-test by Sales Quartiles
Comparing Pre-FAS 154 Earnings versus Post-FAS 154 Earnings

Panel A - Loss Firms:					
	Sales Quartile	**Sales Range (in millions)**	**t Value**	**DF**	**Pr > |t|**
	1st	$0.008 - 121.890	-4.9071	558	0.0000
	2nd	$121.890 - 608.498	-8.1140	558	0.0000
	3rd	$608.498 – 2,912.209	-6.1538	544	0.0000
	4th	$2,912.209 – 213,199.000	-2.8220	499	0.0050
Panel B - Gain Firms:					
	Sales Quartile	**Sales Range (in millions)**	**t Value**	**DF**	**Pr > |t|**
	1st	$0.008 - 121.890	4.0407	93	0.0001
	2nd	$121.890 - 608.498	4.0624	93	0.0001
	3rd	$608.498 – 2,912.209	3.6211	107	0.0005
	4th	$2,912.209 – 213,199.000	4.6835	152	0.0000

Finally, Table 9 presents results by sales quartiles, which were established earlier in Table 2. The results are significant for all quartiles for both sets of firms indicating that the size of the firm was a determining factor in whether FAS 154 would have a statistically significant on earnings.

The results in tables 7, 8 and 9 indicate that in most instances where a sufficient number of observations are present the impact of FAS 154 on earnings would have been significant. These results suggest evidence that FAS 154 will have a significant impact in the future on the reported earnings for the year that a firm makes a change in accounting principle. This is due to the fact that a cumulative-effect on prior years' earnings loss or gain will no longer be reported in earnings in the year a firm makes a change in accounting principle. Instead, retrospective application will be used which does not reveal the cumulative-effect on the income statement. As required by FAS 154 a cumulative-effect loss or gain as of the earliest year reported in the annual report will be reported in retained earnings and thus excluded from earnings.

SUMMARY

The results of this study provide evidence to suggest that if the new accounting standard, FAS 154, had been applied from 1998 through 2004, it would have affected reported earnings. In addition, the mean and median and 75th percentile of the percentage change in earnings would have shown a double-digit impact in some years, industries and size of firms. Finally, evidence exists that the overall impact would have been statistically significant for most years and industries and for all sales quartiles.

REFERENCES

Abdel-khalik, A. R. & McKeown, J. C. (1978). Understanding accounting changes in an efficient market: Evidence of differential reaction. *The Accounting Review*, October, 851- 868.

American Institute of Certified Public Accountants. (1971). *Accounting Principles Board (APB) Opinion No. 20, Accounting Changes*, July, New York, NY.

American Institute of Certified Public Accountants. (2003). *Accounting Trends & Techniques*. New York, NY.

Biddle, G. C. & Lindahl, R. W. (1982). Stock price reactions to lifo adoptions: The association between excess returns and LIFO tax savings. *Journal of Accounting Research*, 20, 3, 551-588.

Binder, J.J. (1983). *Measuring the effects of regulation with stock price data: a new methodology*. Dissertation: The University of Chicago.

Brown, S. & Warner, J. (1980). Measuring security price performance. *Journal of Financial Economics*, 8, 205-258.

Chung, K.Y., Park, T. & Byung, TR. (1996). Differential market reactions to accounting changes away from versus towards common accounting practices. *Journal of Accounting and Public Policy*, Spring, 15, 1; 29-54.

Cready, W.M., Shank, J.K., & McKeown, J.C. (1987). Understanding accounting changes in an efficient market -- a comment, replication, and re-interpretation/analysis of variance issues. *The Accounting Review*, 62 (3) 589-600.

Cushing, B.E. (1974). Accounting changes: The impact of APB No. 20. *Journal of Accountancy*, November, 54-62.

Financial Accounting Standards Board. (2001). *Statement of Financial Accounting Standards No. 143, Accounting for Asset Retirement Obligations.*

Financial Accounting Standards Board. (2005a). *Statement of Financial Accounting Standards No. 154, Accounting Changes and Error Corrections—a replacement of APB Opinion No. 20 and FASB Statement No. 3*.

Financial Accounting Standards Board. (2005b). http://www.fasb.org/project/short-term_intl_convergence.shtml.

Harrison, T. (1977). Different market reactions to discretionary and nondiscretionary accounting changes. *Journal of Accounting Research*, Spring, 84-107.

Holthausen, R. W. & Verrecchia, R. E. (1990). The effect of informedness and consensus on price and volume behavior. *The Accounting Review*, 65 (1), 191–208.

Holthausen, R.W. & Leftwich, R.W. (1983). The economic consequences of accounting choice: implication of costly contracting and monitoring. *Journal of Accounting and Economics*, July, 77-117.

Hughes, J.S., & Ricks, W. (1984). Accounting for retail land sales: Analysis of a mandated change. *Journal of Accounting and Economics*, 6, 101-32.

Johnson, W. B. & Dhaliwal, D.S. (1988). LIFO abandonment. *Journal of Accounting Research*, Chicago: Autumn, 26, 2; 236-72.

Lev, B. (1979). The impact of accounting regulation on the stock market: the case of oil and gas companies. *The Accounting Review*, 54, July, 485-503.

May, G.S. & D.K. Schneider. (1988). Reporting Accounting changes: Are Stricter Guidelines Needed? *Accounting Horizons*, September, 1988, 68-74.

Moore, P., Atkinson, M. & Nix, K. (2003). Cumulative effect of a change in accounting principle: remove it from the income statement. *The CPA Journal*, January, 73, 1, 52-3.

Morse, D. & Richardson, G. (1983). The LIFO/FIFO decision. *Journal of Accounting Research*, Spring, 21, 106-127.

Pfeiffer, G.M. (1980). *Accounting for long term financial commitments: An investigation into the economic effects of lease accounting changes*. Dissertation: Cornell University.

Ricks,W.E. & Hughes, J.S. (1985). Market reactions to a non-discretionary accounting change: the case of long-term investments. *The Accounting Review*, 60, 33–52.

Schipper, K. & Thompson, R. (1983). Evidence on the capitalized value of merger activity for acquiring firms. *Journal of Financial Economics*, 11, 437-467.

Smith, A.J. (1981). *An empirical investigation of the information effects of a change in the financial reporting standards for oil and gas producers: the proposed elimination and subsequent retention of full cost accounting*. Dissertation: Cornell University.

Verrecchia, R.E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica*; November; 50, 6, 1415-31.

AN EMPIRICAL INVESTIGATION OF THE RELATION BETWEEN RELATIVE PERFORMANCE EVALUATION AND R&D EXPENDITURES

Gloria McVay, Winona State University
Myrtle W. Clark, University of Kentucky
Sung Wook Yoon, California State University, Northridge

ABSTRACT

This study extends prior research on relative performance evaluation (RPE) in executive compensation schemes by including a variable for management project choice as posited by Dye (1992). Dye predicts that firms with a large number of project choices and firms with relatively few project choices will use RPE more frequently than other firms will. We use the magnitude of firm's R&D expenditures as a proxy measure of the number of project choices. The results in this study are generally consistent with predictions based on the Dye model. Measures of company stock return to industry or market-wide measures provide evidence of RPE filtering for those companies with relatively few project choices, measured by small R&D expenditures, as well as those with a relatively large number of project choices, measured by large R&D expenditures. For companies with moderate levels of project choices, the industry measures and the market measures of RPE using stock returns are relatively less significant than the other two groups. The results using accounting based measures are inconsistent with the implications of RPE theory and companies do not appear to filter out industry or economy wide shocks when evaluating firm performance.

INTRODUCTION

Agency theory predicts that CEO compensation is determined taking into consideration not only company performance but also the performance of the company relative to the performance of other companies in the same industry and other companies in general (Lazear & Rosen, 1981; Holmstrom 1979, 1982). The underlying premise for this prediction is that company performance is affected by common shocks experienced by all companies in an industry or in the macro-economy; hence, exclusion of common shocks from performance measures provides a better measure for the evaluation of CEO performance. If so, the efficiency of CEO contracts can be improved by filtering out industry and macro-economic effects. Empirical tests of the existence of relative performance evaluation (RPE) in executive compensation contracts have added performance outputs of other firms in the industry and economy-wide effects as explanatory variables. However, these studies provide inconclusive results. For example, Antle and Smith (1986) and Janakiraman,

Lambert, and Larcker (1992) found limited evidence that managerial compensation is adjusted for industry effects. Gibbons and Murphy (1990) found some adjustment in CEO compensation for industry performance, and much more adjustment for changes in market-wide rates of return. Conyon and Gregg (1994) observed no relationship between CEO compensation and industry returns. Studies supporting the implicit use of RPE include Murphy (1985), Antle and Smith (1986), Gibbons and Murphy (1990), Janakiraman, Lambert, and Larcker (1992), Conyon and Gregg (1994), and Bannister and Newman (2003). Studies rejecting the implicit use of RPE include Barro and Barro (1990), Jensen and Murphy (1990), Garen (1994), and Aggarwal and Samwick (1999a and 1999b).

Dye (1992) developed an analytical model of RPE which suggests that prior research failed to document RPE filtering of industry effects because they did not take into consideration the potential impact of investment opportunity choice. According to Dye, the number of project choices the CEO has to choose from when making investment decisions has implications for the relative advantages of using RPE versus absolute performance evaluation. The Dye analytical results imply that when project choice is either eliminated or sufficiently large, the problem of moral hazard found in the principal-agent relationship is minimized. This study extends prior empirical studies of CEO compensation and RPE by considering this project choice variable. Specifically, we test for the presence of RPE in the compensation contracts of CEOs when firms have been partitioned based on the project choice available to the CEOs.

PRIOR RESEARCH

According to RPE theory, rewarding top-level executives based on performance measured relative to aggregate industry performance creates incentives to take actions increasing shareholder wealth while insuring executives against market and industry shocks. Aggregations of individual performance measures (such as industry averages) may be sufficient bases for relative performance schemes (Antle & Smith, 1986). Albuquerque (2006a) argues that firm size should be considered when RPE peer groups are identified. Holmstrom (1979) establishes that the informativeness criterion determines the value of including measures of other agents' performance in the evaluation of the performance of any given agent. Although RPE offers benefits when an important source of common risk or uncertainty exists, it also entails costs (Gibbons & Murphy, 1990). Rewarding performance of workers relative to others may induce costly collusive and shirking behaviors. These types of costs, however, should be trivial when applying RPE to filter shocks common to the industry or the economy to evaluate and compensate CEOs. Hence, using RPE for CEO compensation schemes should provide a linkage between executive compensation and changes in shareholder wealth (Nalbantian, 1993).

Holmstrom (1982) puts forth three propositions concerning RPE: (1) an agent's contract will depend on correlated performance measures of other firms; (2) an average of the correlated performance measures of other firms will be a sufficient summary of the relative information for the

agent's contract; and (3) the common uncertainty shared with other firms will be completely filtered out of the agent's contract. This research relies on these propositions in testing for the presence of RPE in CEO compensation contracts.

Prior empirical research on the use of RPE in CEO compensation contracts provides mixed results. Antle and Smith (1986) found weak support for the use of RPE in the total compensation contracts of 16 of 39 firms in the chemical, aerospace, and electronics industries from 1947 to 1977. Gibbons and Murphy (1990) analyzed 7,757 cash compensation observations from 1,049 firms from 1974 to 1986 and found some adjustment in CEO compensation for industry performance. However, when Gibbons and Murphy analyzed firm and peer-group accounting performance (ROA), their results are inconsistent with the implications of RPE theory. Janikiraman et al. (1992), using market returns, found little empirical support for the implication of the standard agency theory model that the market and industry components of firm performance should be completely removed in determining CEO compensation. In addition, Conyon and Gregg (1994) observed no relationship between executive compensation and industry returns to shareholders and Janikiraman et al. document no support for RPE using financial accounting measures of industry or market-wide return on assets. Albuquerque (2006b) found empirical evidence showing that the use of RPE is negatively associated with a firm's level of growth options.

With regards to the association between project choice (or project selection) and executive compensation, Lambert (1986) examined the incentives of executives to adopt risky projects and showed that the executive and the principal will not always agree regarding which project is best due to the conflict of interest from the agency problem. Barron and Waddell (2003) examined a project selection framework related to compensation that can influence the executive's effort choice. They argued that the effect of firm characteristics, such as R&D expenditures and the market-to-book ratio, on the structure of executive compensation seem to be important in interpreting the relative informativeness of performance signals.

RPE AND PROJECT CHOICE

Dye (1992) presents an analytical model of RPE and project selection. The model predicts that the number of projects available to a manager is related to the effectiveness of RPE in minimizing the problem of moral hazard between principal and agent. The results demonstrate that gains achievable from RPE vary nonmonotonically with the number of industries or projects to which the agent has access. Dye (1992) shows that when an agent has limited discretion in project choice, RPE yields little improvement in contracting. However, as the number of project choices expands to a sufficiently large level, RPE is found to be effective in reducing the moral hazard problem between principal and agent. The greatest reductions in the expected cost of compensating the agent via RPE occur when the agent's discretion in project choice is eliminated. While these results must be interpreted with caution and considered in the context of the underlying assumptions and constraints of the model, they provide a testable implication of RPE theory. Dye (1992) suggests

that prior empirical studies were marginally successful in identifying the association between RPE and executive compensation due to underspecification of the models employed. Specifically, he suggests that an important extension of his analytical work is the empirical testing of a refined model that includes a previously omitted variable – the number of project choices available to the agent (CEO). This study is designed to include this variable, thereby extending prior empirical studies and testing the analytical model developed by Dye (1992).

RESEARCH METHODOLOGY

This research is designed to investigate empirically whether CEOs are compensated as if their performance is evaluated relative to the performance of competitors. In this study, the test for the presence of RPE is refined by partitioning firms based upon project choice available, measured by magnitude of firm's R&D expenditures, to the CEO.

Executive Compensation

RPE theory predicts that CEO compensation will vary directly with firm performance, after filtering out industry and macro-economic shocks. CEO compensation typically comprises a portfolio of incentive vehicles, including a bonus plan, executive stock options, as well as other long-term performance incentives. Consistent with previous studies (e.g., Baber, Janakiraman, & Kang, 1996; Antle & Smith, 1986), we examine the CEO's total compensation package. Our dependent variable is the change in log value of total annual compensation, which includes salary, bonus, other annual compensation, total value of restricted stock granted, total value of stock options granted, and long-term payouts.

Sample Selection

The companies selected for study were included in Standard & Poor's ExecuComp database. Executive compensation for the period of 1996–2004 was collected for firms meeting the following selection criteria:

1. For a given year under study, the executive must have been the CEO of the company for at least two fiscal years. This criterion minimizes the potential for the inclusion of signing bonuses in compensation. It also ensures that the base for measuring percentage change measures a full year's compensation.

2. Data were excluded for a given year when company performance measures or the data for project choice were unavailable on ExecuComp or Compustat.

This procedure provided a sample containing 2,738 observations across 212 4-digit SIC codes and 50 2-digit SIC codes.

Partitioning the Sample

According to Dye (1992), companies with relatively few or a significant number of project choices are more likely to use RPE to compensate their CEOs. Following prior research (Barron and Waddell, 2003), Research and Development expenditures deflated by the book value of total assets serves as a proxy for the number of project choices available to a CEO. The definition of research and development expenditure by the Statement of Financial Accounting Standard (SFAS) No. 2 suggests that the magnitude of R&D expenditures may serve as an indicator for executives' new product development activities.[1] Use of this proxy implies that in the context of RPE, modified to take into consideration Dye's analytical results, the CEO has responsibility for the allocation of company resources to the available project choices. In any particular year, the allocation of company resources to the available project choices is an important determinant of the company's success and the CEO's performance. A test of the average effects for the entire sample allows inferences about the use of RPE in general. This technique was employed in prior empirical research with limited findings (Antle & Smith, 1986; Gibbons & Murphy, 1990), and is replicated here for purposes of comparison. The primary focus of this study is to determine if there is a relationship between the magnitude of project choices available to a CEO and the firm's use of RPE. The sample is partitioned into five quintile groups based on the magnitude of R&D expenditures to test for differences in the usage of RPE to compensate CEOs with more or less extensive investment opportunities. To make the difference between partitioned groups clear, the 2nd quintile and the 4th quintile have been eliminated from the sample and only the 1st quintile, the 3rd quintile, and the 5th quintile are examined. In this study, the 1st quintile group represents firms with a small magnitude of R&D expenditures, the 5th quintile group represents firms with a large magnitude of R&D expenditures, and 3rd quintile group represents firms with a moderated magnitude of R&D expenditures

Hypotheses

The research question addressed in this study is whether the use of RPE in compensation schemes for CEOs will vary nonmonotonically with the magnitude of R&D expenditures. According to Dye (1992), RPE is more effective in reducing the moral hazard of contracting when the number of choices available to the agent (CEO) is very limited or sufficiently extensive. Therefore, we hypothesize that firms with relatively small or substantially sizable magnitude of R&D expenditures will observe the use of RPE more frequently. Conversely, firms with a moderate magnitude of R&D expenditures will employ RPE less frequently. Specifically, the following hypotheses are tested in this study:

Hypothesis 1: Firms with a small magnitude of R&D expenditures will make more frequent use of RPE than firms with a moderate magnitude of R&D expenditures will.

Hypothesis 2: Firms with a large magnitude of R&D expenditures will make more frequent use of RPE than firms with a moderate magnitude of R&D expenditures will.

Measuring Performance

According to RPE theory, CEO compensation should reflect the absolute performance measure for the company and an adjustment for industry and/or macro-economic shocks. Absolute performance measures may be based on book values or market values. We utilize two performance measures: return on assets (ROA), and the market return to common stockholders (RET). Both measures were tested in prior research.

ROA and RET were collected for each company in the sample and for all companies for which Compustat reports these returns. Compustat calculates RET and ROA as follows:

$$RET_{it} = 100 \times (((price_{it} + dividends_{it} + cash\ equivalents_{it})/price_{it-1}) - 1)$$

$$ROA_{it} = 100 \times ((Income\ before\ extraordinary\ items_{it})/average\ assets_{it})$$

For this study, these measures were divided by 100 and then 1 was added to produce a relative well-offness measure similar to a price relative. Such a measure minimizes the effect of misinterpreting the relative performance of a company when the company return is positive and the industry or market-wide return is negative. Previous investigations of RPE (e.g., Antle & Smith, 1986; Gibbons & Murphy, 1990; and Janakiraman, Lambert, and Larcker (1992) utilized RET and/or ROA measures without considering this potential confounding effect. As a result, they may not have measured relative performance for some of the companies in their sample. For example, if a company and the industry experience a loss and the company's loss is large relative to the industry, the company's loss relative to the industry loss would produce an appearance that the company is outperforming the industry, when in fact the company is underperforming. Similarly, if the company has a positive return and the industry has a negative return, use of unadjusted ROA and RET measures would provide the appearance of underperformance. It is therefore possible that the weak results of prior investigations of RPE can be at least partially explained by the measurement of the explanatory variables.

Measuring Relative Performance

If RPE is present, the absolute performance measure will be evaluated relative to industry and/or macro-economic performance measures. The macro-economic measures utilized in this study are the average ROA and the annual market return (adjusted as described above) for all companies in the Compustat database for which Compustat reports ROA and RET.

The RPE theory implies that firm performance is evaluated relative to the performance of companies of similar risk (Holmstrom, 1982). Firms facing similar risks and uncertainties should be affected similarly. Hence, the difference in performance of an individual manager and that of an appropriately weighted peer group conveys valuable information regarding that manager's performance (Nalbantian, 1993). Logically, companies in the same industry face similar risks. However, companies within a given industry classification may display differences in risk-return measures. Hence, an appropriate comparison of company performance relative to a peer group index would give more weight to the returns of those companies with risk measures that are more closely correlated with the company being evaluated. Antle and Smith (1986) employed an industry index weighting scheme utilizing correlation coefficients between the sample firm's return and those of other firms in the same industry. The resulting index gave more weight to those returns that were more closely correlated with the sample firm. The Antle and Smith results are based on time-series regressions for the individual companies in their sample. The current study utilizes cross-sectional regression and weights peer group returns using betas reported by Compustat. Betas provide information on the relative riskiness of companies. Similar betas for companies in the same industry would indicate that the returns of those companies are correlated with each other because they are correlated in the same way with the market return.

For this study, the industry return utilized for a given sample company for a given test year is calculated using a weighted average that weights companies' returns that are similar to the sample firm more heavily. We calculate weights utilizing company betas. The weight, w_{it}, assigned to firm *i* in the construction of the industry ROA for all companies in the same two-digit industry SIC code as sample firm s is given as:

$$w_{it} = 1 - |b_{st} - b_{it}| / \text{å} b_{jt}$$

where b_{st} = the beta for sample firm s for year t
b_{it} = the beta for non-sample firm i for year t
j indicates all firms in the industry except for the sample firm

An industry ROA is then constructed for each sample firm as follows:

$$ROA_{It} = \text{å}\, w_{jt} \times ROA_{jt}$$

Similarly, an industry market return, RET, is also calculated for each sample firm as:

$$RET_{It} = \mathring{a}\ w_{jt} \times RET_{jt}$$

To provide additional testing, the above calculations were repeated for each sample company using four-digit SIC codes. If the absolute performance measure is greater relative to the industry or macro-economic performance measure, the CEO would be perceived as performing well relative to others in the industry or the economy as a whole. Alternatively, poor performance relative to others would indicate that the CEO's relative performance is low. We measure relative performance as the ratio of absolute performance to the industry/macro-economic benchmark minus one. That is, the measurement for a sample company's RPE measure for ROA relative to that of the industry for year *t*, $IROARPE_{st}$, and for the company's RPE measure for RET relative to the industry market return, $IRETRPE_{st,}$ are calculated as follows:

$$IROARPE_{st} = ROA_{st} / ROA_{It} - 1 IRETRPE_{st} = RET_{st} / RET_{It} - 1$$

Good relative performance would be positive. Poor relative performance would be negative. Hence, there should be a positive correlation between CEO compensation and the RPE measure utilized. We calculate similar RPE measures for company performance relative to the macro-economic ROA, $MROARPE_{st}$, and to the macro-economic RET, $MROARPE_{st}$ as follows:

$$MROARPE_{st} = ROA_{st} / ROA_{It} - 1 MRETRPE_{st} = RET_{st} / RET_{It} - 1$$

Modeling Relative Performance and CEO Compensation

If CEO compensation reflects RPE, the change in log value of compensation from one year to the next should be positively correlated with company performance relative to industry and/market-wide performance. We estimate the following regression equations for our total sample of firms and for sub-samples based on the number of product lines.

$$Log(DCOMP)_{st} = a_s + g_s IROARPE_{st} + f_s IRETRPE_{st} + \lambda_s YearDummy + e_{st}$$
$$Log(DCOMP)_{st} = a_s + g_s MROARPE_{st} + f_s MRETRPE_{st} + \lambda_s YearDummy + e_{st}$$

A significant positive g_s regression coefficient would be consistent with RPE of ROA for CEO performance relative to other companies in the same industry or macro-economic performance. Similarly, a significant positive f_s would indicate RPE of return for CEO performance relative to other companies in the same industry or macro-economic performance. By including a "year dummy" variable, the problem arising from panel data sample can be mitigated.

RESULTS

Descriptive Statistics

Table 1 reports descriptive statistics for the RPE sample companies. Note that these companies display similar ROA ratios across the three sample partitions. Moreover, the mean market returns between partitions are not significantly different from each other. Thus, absent evaluation of performance relative to industry or economy-wide shocks, partitioning by investment choice is unlikely to be associated with differences in the way companies reward CEO performance.

Table 1: Descriptive Sample Statistics

Attribute	Lowest	Moderate	Highest	Total
	(n=547)	(n=548)	(n=547)	(n=2,738)
Change in log of CEOs' Total Compensation				
Mean	0.060794	0.078530	0.048537	0.068631
Median	0.055520	0.080228	0.069345	0.077008
Standard Deviation	0.388600	0.334860	0.387290	0.357170
R&D Expenditures (in thousands)				
Mean	1.602349	93.47184	828.9394	237.7386
Median	0.000000	38.92600	244.5420	33.18000
Standard Deviation	19.10261	204.5814	1399.000	733.9887
Total Assets (in thousands)				
Mean	6262.135	8613.943	14634.69	9596.336
Median	2111.660	1894.037	4115.700	2422.050
Standard Deviation	15898.00	41516.00	42652.00	35220.00
Return on Assets				
Mean	0.054452	0.054134	0.055482	0.054596
Median	0.050100	0.056330	0.059760	0.055705
Standard Deviation	0.049934	0.049880	0.084070	0.061470
Stock Returns				
Mean	0.150895	0.145049	0.135678	0.139721
Median	0.098920	0.128675	0.111000	0.121225
Standard Deviation	0.422790	0.354130	0.410500	0.380490

Aggregate Results

Regression results for all 2,738 observations are presented in Table 2. The findings are consistent with the results of prior studies. Panel A shows that RPE filtering is apparent for the economy wide measure of market-based returns. The coefficient is positive and statistically significant. A weak relationship is observed between CEO compensation and market-based industry measures of relative performance. Panel B shows that the model containing only financial accounting measures displays insignificant coefficients for both the industry variable and the market-wide variable. Moreover, Panel C demonstrates that RPE measures using market- based returns are associated with CEO compensation, but accounting based RPE measures do not appear to be related to CEO compensation, which is consistent with the finding of Gibbons and Murphy (1990).

Table 2: Regression Coefficients for Aggregate RPE Sample

(Dependent Variable=Change in the log value of CEOs' Total Compensation, Sample size = 2,738)

(Panel A) Aggregate Sample using Market-Based RPE Measures					
Intercept	**g_s**	**t-value**	**f_s**	**t-value**	**Adj. R^2**
RPE Measures based on 4-digit SIC Codes and Market Index					
	IRETRPE4		**MRETRPE**		
0.08519	0.01513	1.65	0.33237	14.52	0.1297
RPE Measures based on 2-digit SIC Codes and Market Index					
	IRETRPE2		**MRETRPE**		
0.08752	0.06563	2.06	0.28796	7.94	0.1302
The results for year dummy variables are not shown					
(Panel B) Aggregate Sample using Accounting-Based RPE Measures					
Intercept	**g_s**	**t-value**	**f_s**	**t-value**	**Adj. R^2**
RPE Measures based on 4-digit SIC Codes and Market Index					
	IROARPE4		**MROARPE**		
0.1789	0.00000	1.17	-0.00024	-1.36	0.0351
RPE Measures based on 2-digit SIC Codes and Market Index					
	IROARPE2		**MROARPE**		
0.1857	-0.00001	-1.64	-0.00024	-1.38	0.0355
The results for year dummy variables are not shown					
(Panel C) Aggregate Sample using both Market-Based RPE Measures and Accounting-Based RPE Measures					
Intercept	**g_s**	**t-value**	**f_s**	**t-value**	**Adj. R^2**

Table 2: Regression Coefficients for Aggregate RPE Sample					
(Dependent Variable=Change in the log value of CEOs' Total Compensation, Sample size = 2,738)					
RPE Measures based on 4-digit SIC Codes					
	IROARPE4		**IRETRPE4**		
0.05439	0.00000	1.13	0.07660	9.10	0.0629
RPE Measures based on 2-digit SIC Codes					
	IROARPE2		**IRETRPE2**		
0.07445	-0.00001	-1.56	0.27508	15.26	0.1109
RPE Measures based on Market Index					
	MROARPE		**MRETRPE**		
0.18811	-0.00027	-1.63	0.35027	17.26	0.1297
The results for year dummy variables are not shown					

Effects of Investment Choice and Combined Measures of Market-based and Accounting-based RPE Measures

Table 3 reports the regression results for the RPE sample partitioned by the level of R&D expenditures. For companies with all three levels of R&D expenditures in Panel A, the coefficients for the industry accounting RPE measure using 4-digit SIC codes are insignificant, while the market return-based relative performance measures are positively correlated with CEO compensation. As hypothesized, the coefficient (0.06006) of IRETRPE4 for the companies with moderate level of R&D expenditures is less than the coefficient (0.07543) for the companies with lowest level of and the coefficient (0.10516) for the companies with highest level of R&D expenditures. As in Panel B and C, results for the regression for the RPE measures using either 2-digit codes or market index are similar to the results in Panel A. Regression coefficients of IRETRPE4 for the companies with moderate level of R&D expenditures are markedly different from the coefficients for firms with lowest or highest level of R&D expenditures.

Table 3: Regression Coefficients for Partitioned RPE Sub-Samples using both Market-Based RPE Measures and Accounting-Based RPE Measures			
(Dependent Variable=Change in the log value of CEOs' Total Compensation)			
(Panel A) RPE Measures based on 4-digit SIC Codes			
Parameters	**R&D Expenditures**		
	Lowest	**Moderate**	**Highest**
Intercept	-0.07001	0.18271	-0.07368
IROARPE4	-0.00000	0.00001	0.00001

Table 3: Regression Coefficients for Partitioned RPE Sub-Samples using both Market-Based RPE Measures and Accounting-Based RPE Measures			
(Dependent Variable=Change in the log value of CEOs' Total Compensation)			
t-value	-0.15	1.16	1.27
IRETRPE4	0.07543	0.06006	0.10516
t-value	4.37	3.46	4.36
Adj. R^2	0.0394	0.0461	0.0666
No. of obs.	547	548	547
The results for year dummy variables are not shown			
(Panel B) RPE Measures based on 2-digit SIC Codes			
Parameters	**R&D Expenditures**		
	Lowest	**Moderate**	**Highest**
Intercept	-0.12158	0.18665	-0.04633
IROARPE2	-0.00000	-0.00000	0.00000
t-value	-0.14	-0.73	0.19
IRETRPE2	0.28988	0.22874	0.32881
t-value	7.85	5.51	7.50
Adj. R^2	0.1079	0.0763	0.1227
No. of obs.	547	548	547
The results for year dummy variables are not shown			
(Panel C) RPE Measures based on Market Index			
Parameters	**R&D Expenditures**		
	Lowest	**Moderate**	**Highest**
Intercept	-0.06448	0.18649	0.23325
MROARPE	0.00013	0.00006	-0.00079
t-value	0.27	0.16	-2.14
MRETRPE	0.43677	0.29256	0.33471
t-value	9.50	6.34	7.49
Adj. R^2	0.1492	0.0906	0.1289
No. of obs.	547	548	547
The results for year dummy variables are not shown			

These regression results are consistent with RPE theory and provide support our hypotheses based on the Dye (1992) model. Companies with relatively few project choices, measured by the

bare minimum amount of R&D expenditures, filter out the effects of industry or economy-wide shocks better when compensating CEOs than companies with moderate project choices do. Companies with limited project choices, measured by moderate spending of R&D expenditures, relatively less react to either industry or economy-wide measures of performance when evaluating executive performance compared to companies with relatively many project selections available.

Market-based Measures and the Effects of Investment Choice

Market-based regression results are similar to the results from the combined measures of accounting-based and market-based RPE (see Table 4). The economy-wide variable, MRETRPE, displays a positive coefficient and is significant at Pr>|t| = <0.001. For sample firms with moderate level of R&D expenditures, explanatory powers of the regression models, measured by adjusted R-squares, are substantially less than the explanatory powers of firms with lowest or highest level of R&D expenditures. These results also provide support for the RPE hypothesis. However, industry market-measures of RPE are not statistically significant in the model using 4-digit SIC codes.

Table 4: Regression Coefficients for Partitioned RPE Sub-Samples using Market-Based RPE Measures

(Dependent Variable=Change in the log value of CEOs' Total Compensation)

(Panel A) RPE Measures based on 4-digit SIC Codes and Market Index

Parameters	**R&D Expenditures**		
	Lowest	**Moderate**	**Highest**
Intercept	-0.02869	0.20077	-0.04999
IRETRPE4	0.01373	0.02196	0.01970
t-value	0.77	1.20	0.72
MRETRPE	0.42123	0.26915	0.31429
t-value	8.35	5.39	5.97
Adj. R^2	0.1500	0.0929	0.1223
No. of obs.	547	548	547

The results for year dummy variables are not shown

(Panel B) RPE Measures based on 2-digit SIC Codes and Market Index

Parameters	**R&D Expenditures**		
	Lowest	**Moderate**	**Highest**
Intercept	-0.05033	0.20854	-0.04508
IRETRPE2	0.08702	0.00051	0.18365
t-value	1.67	0.01	1.82
MRETRPE	0.35703	0.29168	0.16481

Table 4: Regression Coefficients for Partitioned RPE Sub-Samples using Market-Based RPE Measures			
(Dependent Variable=Change in the log value of CEOs' Total Compensation)			
t-value	5.37	2.99	1.60
Adj. R^2	0.1535	0.0905	0.1268
No. of obs.	547	548	547
The results for year dummy variables are not shown			

Accounting-based Measures and the Effects of Investment Choice

Financial accounting-based regression results are quite different from the market-based results (see Table 5). In the regression model using 4-digit SIC codes or 2-digit SIC codes, the industry accounting-measures of RPE are not significant. Similar to finding of Gibbons and Murphy (1990), accounting-based measures do not seem to filter out the effects of industry or economy-wide shocks when compensating CEOs.

Table 5: Regression Coefficients for Partitioned RPE Sub-Samples using Accounting-Based RPE Measures			
(Dependent Variable=Change in the log value of CEOs' Total Compensation)			
(Panel A) RPE Measures based on 4-digit SIC Codes and Market Index			
Parameters	**R&D Expenditures**		
	Lowest	**Moderate**	**Highest**
Intercept	-0.15542	0.22452	0.23343
IROARPE4	-0.00000	0.00001	0.00001
t-value	-0.32	1.15	1.35
MROARPE	0.00035	-0.00006	-0.00078
t-value	0.69	-0.15	-2.01
Adj. R^2	0.006	0.0248	0.0408
No. of obs.	547	548	547
The results for year dummy variables are not shown			
(Panel B) Accounting-Based RPE Measures based on 2-digit SIC Codes			
Parameters	**R&D Expenditures**		
	Lowest	**Moderate**	**Highest**
Intercept	-0.15944	0.23064	0.23028
IROARPE2	-0.00000	-0.00001	0.00000
t-value	-0.31	-0.96	0.23
MROARPE	0.00034	-0.00007	-0.00077

Table 5: Regression Coefficients for Partitioned RPE Sub-Samples using Accounting-Based RPE Measures			
(Dependent Variable=Change in the log value of CEOs' Total Compensation)			
t-value	0.69	-0.17	-1.99
Adj. R^2	0.006	0.0241	0.0376
No. of obs.	547	548	547
The results for year dummy variables are not shown			

SUMMARY AND CONCLUSIONS

This study provides evidence supporting Dye's (1992) analytical model of relative performance evaluation and project choice. Dye predicts that firms with a large number of project choices and firms with relatively few project choices will use RPE more frequently than other firms will. The regression results in this study are generally consistent with predictions based on the Dye model. Measures of company stock return to industry or market-wide measures provide evidence of RPE filtering for those companies with relatively few project choices, measured by small R&D expenditures, as well as those with a relatively large number of project choices, measured by large R&D expenditures. For companies with moderate levels of project choices, the industry measures and the market measures of RPE using stock returns are relatively less significant. Regression results using accounting based measures are inconsistent with the implications of RPE theory and companies do not appear to filter out industry or economy wide shocks when evaluating firm performance.This study is important because empirical findings that support the predictions of the Dye (1992) analytical model suggest that the costs of implementing and maintaining the RPE-based compensation scheme are not warranted in some cases but are cost-efficient in others. Moreover, significant findings consistent with the predictions of Dye's model help explain the weak and inconsistent findings of prior empirical studies that omitted the number of investment choices as an explanatory variable.

ENDNOTES

[1] According to SFAS No. 2, *Research* is "planned search or critical investigation aimed at discovery of new knowledge with the hope that such knowledge will be useful in developing a new product or service or a new process or technique or in bringing about a significant improvement to an existing product or process." *Development* is "the translation of research findings or other knowledge into a plan or design for a new product or process or for a significant improvement to an existing product or process whether intended for sale or use."

REFERENCES

Albuquerque, A. (2006a). Who Are Your Peers?: A Study of Relative Performance Evaluation. Working Paper, Boston University

Albuquerque, A. (2006b). Do Growth Options Firms Use Less Relative Performance Evaluation? Working Paper, Boston University

Aggarwal, R. & Samwick, A. (1999a). Executive compensation, strategic compensation, and relative performance evaluation: Theory and evidence. *Journal of Finance,* 54,1999-2043.

Aggarwal, R. & Samwick, A. (1999b). The other side of the tradeoff: The impact of risk on executive compensation. *Journal of Political Economy,* 107, 65-105.

Antle, R., & Smith, A. (1986). An empirical investigation of the relative evaluation of corporate executives. *Journal of Accounting Research,* 24(1), 1-39.

Baber, W. R., Janakiraman, S. N., & Kang, S-H. (1996). Investment opportunities and the structure of executive compensation. *Journal of Accounting and Economics,* 21, 297-318.

Bannister, J. W., & Newman, H. A. (2003). Analysis of corporate disclosures on relative performance evaluation. *Accounting Horizons,* 17, 235-246.

Barro, J. R., & Barro, R. J. (1990). Pay, performance, and turnover of bank CEOs. *Journal of Labor Economics,* 8, 448-481.Barron, J. M, & Waddell, G. R. (2003), Executive rank, pay, and project selection. *Journal of Financial Economics*, 67, 305-349

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy,* 81(5), 637-654.

Conyon, M. J., & Gregg, P. (1994). Pay at the top: A study of the sensitivity of top director remuneration to company specific shocks. *National Institute Economic Review,* 149, 83-93.

Dye, R. A. (1992). Relative performance evaluation and project selection. *Journal of Accounting Research,* 30(1), 27-52.

Garen, J. (1994). Executive compensation and principal-agent theory. *Journal of Political Economy,* 102, 1175-1199.

Gibbons, R. & Murphy, K. J. (1990). Relative performance evaluation for chief executive officers. *Industrial and Labor Relations Review,* 43 (Special Issue), 30-51.

Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics,* 10, 74-91.

Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics,* 13, 324-340.

Janakiraman, S., Lambert, R., & Larcker, D. (1992). An empirical investigation of the relative performance evaluation hypothesis. *Journal of Accounting Research,* 30(1), 53-69.

Jenson, M. C., & Murphy, K. J. (1990). Performance pay and top-management incentives. *Journal of Political Economy,* 98, 225-264.Lambert, R. A. (1986). Executive effort and the selection of risky projects. *Rand Journal of Economics*, 17, 77-88

Lazear, E., & Rosen, S. (1981). Rank-ordered tournaments as optimum incentive labor contracts. *Journal of Accounting Research,* 21(1), 177-191.

Murphy, K. J. (1985). Corporate performance and managerial remuneration: An empirical examination. *Journal of Accounting and Economics,* 7, 11-42.

Nalbantian, H. R. (1993). Performance indexing in stock option and other incentive compensation programs. *Compensation and Benefits Review,* 25(5), 25-41.

Allied Academies

invites you to check our website at

www.alliedacademies.org

for information concerning

conferences and submission instructions