ADVANCES IN MANAGEMENT ACCOUNTING

VOLUME 14

ADVANCES IN MANAGEMENT ACCOUNTING

MARC J. EPSTEIN
JOHN Y. LEE
Editors

ADVANCES IN MANAGEMENT ACCOUNTING

ADVANCES IN MANAGEMENT ACCOUNTING

Series Editors: Marc J. Epstein and John Y. Lee

Volume 1–13: Advances in Management Accounting

ADVANCES IN MANAGEMENT ACCOUNTING VOLUME 14

ADVANCES IN MANAGEMENT ACCOUNTING

EDITED BY

MARC J. EPSTEIN
Harvard University and Rice University, USA

JOHN Y. LEE
Pace University, USA

ELSEVIER
JAI

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

ELSEVIER B.V.
Radarweg 29
P.O. Box 211
1000 AE Amsterdam,
The Netherlands

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

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

ELSEVIER Ltd
84 Theobalds Road
London
WC1X 8RR
UK

First edition 2005

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

ISBN-10: 0-7623-1243-2
ISBN-13: 978-0-7623-1243-6
ISSN: 1474-7871 (Series)

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

CONTENTS

This page intentionally left blank

LIST OF CONTRIBUTORS

Robert M. Brown	Pamplin College of Business, Virginia Polytechnic Institute and State University, VA, USA
Stan Davis	Babcock Graduate School of Management, Wake Forest University, NC, USA
Gerald K. DeBusk	John A. Walker College of Business, Appalachian State University, NC, USA
Alan S. Dunk	School of Business & Government, University of Canberra, Australia
Marc J. Epstein	Jones Graduate School of Management, Rice University, TX, USA and Harvard Business School, MA, USA
Kenneth J. Euske	Graduate School of Business and Public Policy, Naval Postgraduate School, CA, USA
John H. Evans, III	Katz Graduate School of Business, University of Pittsburgh, PA, USA
Eric G. Flamholtz	Anderson School of Management, University of California, Los Angeles, CA, USA
Robert Greenberg	College of Business and Economics, Washington State University, WA, USA
Wayne A. Johnson	College of Business Administration, Bowling Green State University, OH, USA
Larry N. Killough	Pamplin College of Business, Virginia Polytechnic Institute and State University, VA, USA
James M. Kohlmeyer, III	East Carolina University, NC, USA

Chao-Hsiung Lee	National Chung Cheng University, Taiwan
John Y. Lee	Lubin School of Business, Pace University, NY, USA
Andrew Leone	Simon Graduate School of Business Administration, University of Rochester, NY, USA
Adam S. Maiga	School of Business Administration, University of Wisconsin-Milwaukee, WI, USA
Mary A. Malina	Graduate School of Business and Public Policy, Naval Postgraduate School, CA, USA
Yasuhiro Monden	Institute of Policy and Planning Sciences, University of Tsukuba, Japan
Nandu J. Nagarajan	Katz Graduate School of Business, University of Pittsburgh, PA, USA
Atieno A. Ndede-Amadi	College of Business Administration, Bowling Green State University, OH, USA
Thomas R. Nunamaker	College of Business and Economics, Washington State University, WA, USA
Jeffrey F. Shields	School of Business, University of Southern Maine, ME, USA
Michael D. Shields	Eli Broad Graduate School of Management, Michigan State University, MI, USA
Ken C. Snead, Jr.	College of Business Administration, Bowling Green State University, OH, USA
Priscilla S. Wisner	College of Business, Montana State University, MT, USA

EDITORIAL BOARD

STATEMENT OF PURPOSE AND REVIEW PROCEDURES

Advances in Management Accounting (*AIMA*) is a professional journal whose purpose is to meet the information needs of both practitioners and academicians. We plan to publish thoughtful, well-developed articles on a variety of current topics in management accounting, broadly defined.

Advances in Management Accounting is to be an annual publication of quality applied research in management accounting. The series will examine areas of management accounting, including performance evaluation systems, accounting for product costs, behavioral impacts on management accounting, and innovations in management accounting. Management accounting includes all systems designed to provide information for management decision making. Research methods will include survey research, field tests, corporate case studies, and modeling. Some speculative articles and survey pieces will be included where appropriate.

AIMA welcomes all comments and encourages articles from both practitioners and academicians.

REVIEW PROCEDURES

AIMA intends to provide authors with timely reviews clearly indicating the acceptance status of their manuscripts. The results of initial reviews normally will be reported to authors within eight weeks from the date the manuscript is received. Once a manuscript is tentatively accepted, the prospects for publication are excellent. The author(s) will be accepted to work with the corresponding Editor, who will act as a liaison between the author(s) and the reviewers to resolve areas of concern. To ensure publication, it is the author's responsibility to make necessary revisions in a timely and satisfactory manner.

EDITORIAL POLICY AND MANUSCRIPT FORM GUIDELINES

1. Manuscripts should be type written and double-spaced on $8^1/_2$" by 11" white paper. Only one side of the paper should be used. Margins should be set to facilitate editing and duplication except as noted:
   a. Tables, figures, and exhibits should appear on a separate page. Each should be numbered and have a title.
   b. Footnote should be presented by citing the author's name and the year of publication in the body of the text; for example, Ferreira (1998); Cooper and Kaplan (1998).
2. Manuscripts should include a cover page that indicates the author's name and affiliation.
3. Manuscripts should include on a separate lead page an abstract not exceeding 200 words. The author's name and affiliation should not appear on the abstract.
4. Topical headings and subheadings should be used. Main headings in the manuscript should be centered, secondary headings should be flush with the left hand margin. (As a guide to usage and style, refer to the William Strunk, Jr., and E.B. White, *The Elements of Style*.)
5. Manuscripts must include a list of references, which contain only those works actually cited. (As a helpful guide in preparing a list of references, refer to Kate L. Turabian, *A Manual for Writers of Term Papers, Theses, and Dissertations*.)
6. In order to be assured of anonymous review, authors should not identify themselves directly or indirectly. Reference to unpublished working papers and dissertations should be avoided. If necessary, authors may indicate that the reference is being withheld for the reason cited above.
7. Manuscripts currently under review by other publications should not be submitted. Complete reports of research presented at a national or regional conference of a professional association and "State of the Art" papers are acceptable.
8. Four copies of each manuscript should be submitted to John Y. Lee at the address below under Guideline 11.
9. A submission fee of $25.00, made payable to Advances in Management Accounting, should be included with all submissions.
10. For additional information regarding the type of manuscripts that are desired, see "*AIMA* Statement of Purpose."

11. Inquires concerning Advances in Management Accounting may be directed to either one of the two editors:

Marc J. Epstein
Jones Graduate School of Administration
Rice University
Houston, TX 77251-1892

John Y. Lee
Lubin School of Business
Pace University
Pleasantville, NY 10570-2799

This page intentionally left blank

INTRODUCTION

This volume of *Advances in Management Accounting (AIMA)* begins with a paper by Evans, Leone, and Nagarajan on non-financial performance measures, or quality-based incentives, in particular, in the healthcare industry. This study examines the economic consequences of non-financial measures of performance in contracts between Health Maintenance Organizations (HMOs) and primary care physicians (PCPs). The authors examine how quality provisions in HMO–PCP contracts affect utilization (patient length of stay in the hospital), patient satisfaction, and HMO costs. In the second paper, Shields and Shields review the research on revenue drivers by reference to five revenue-driver models in the accounting literature. The revenue drivers identified by quantitative empirical research are located in a revenue-driver model based on their levels of analysis (customer, product, organization, and industry) and other characteristics of a revenue driver–revenue relation.

The next paper by DeBusk, Killough, and Brown examines potential cognitive difficulties inherent in the use of performance-measurement systems. They examine the potential for emphasizing financial measures as compared to non-financial measures in the evaluation of an organization's overall performance. The results suggest users of performance-measurement data will emphasize historical financial measures. Alan Dunk's paper follows with a discussion of the quality of information system information, corporate environmental integration, product innovation, and product quality to investigate the extent to which these variables influence financial and non-financial performance. All four independent variables were found to enhance performance assessed in non-financial terms. In contrast, the results show that product innovation alone influences financial performance. Dunk suggests that the efficacy of these factors may be more effectively assessed by evaluating their impact on performance measured in non-financial terms and the inclusion of non-financial measures in performance-evaluation models should enhance control system functioning.

The paper by Epstein and Wisner examines the relationship between management control systems and structures and environmental compliance. Using data from 236 Mexican manufacturing facilities, they test the

applicability of management control theory in Mexican industry. They argue that success in compliance with environmental regulations is significantly associated with the degree of management commitment, planning, belief systems, measurement systems, and rewards. This study contributes evidence about the implementation of environmental strategies in organizations.

The next paper by Eric Flamholtz examines the implications for accounting, information, and control of a growing body of research to develop and empirically test a holistic model of organizational success and failure in entrepreneurial organizations at different stages of growth. The initial model proposes that there are six key factors or "strategic building blocks" of successful organizations, and the six key variables must be designed as a holistic system, which has been termed "The Pyramid of Organizational Development." In the next paper, Euske and Malina SEQ CHAPTER comment on how to improve and build upon this Pyramid model with an eye to the more general question of what we should expect of performance measurement models. Their discussion includes model characteristics, model testing, and then implications for such models. Flamholtz, in his reply to Euske and Malina, states that they have presented a thoughtful and constructive critique of his article but he disagrees with some of the questions and criticisms they have raised.

The paper by Lee, Lee, and Monden examines the link between product development organization and target cost management. They investigate the interactive effects of alternative product development organizations, methods for setting target costs, and alternative decision-making authority in assigning targets. Using a questionnaire survey of Japanese manufacturers, the authors provide some early evidence on those interactive effects. Adam Maiga's study uses structural equation modeling to investigate the relationships between environmental uncertainty, budget communication, budget influence, budget goal commitment, and managerial performance. Based on the study of 173 U.S. individual managers, he shows that environmental uncertainty significantly affects both budget communication and budget influence, which in turn, impact budget goal commitment.

In the next paper, Davis and Kohlmeyer report on their examination of the effect of the employee rank on attitudes and performance when supervisors establish budgeted standards of performance. This paper considers a variable (employee rank) not considered in prior related studies. They report that the impact of attitudes on performance is moderated by the rank of the employee within the organization. The paper by Snead, Johnson and Ndede-Amadi attempts to determine if expectancy theory would be useful in

explaining the motivation of managers to incorporate activity-based costing information into their job. Data obtained from two experiments employing a judgment modeling methodology support the relevance of both the valence and force models of expectancy theory. Next, Greenberg and Nunamaker examine the possible problems of using input–output models when outputs are difficult to quantify within an agency theory perspective and illustrate the potential problems using recent proposals in the U.K. for evaluating and rewarding police unit performance.

We believe the 13 articles in Volume 14 represent relevant, theoretically sound, and practical studies the discipline can greatly benefit from. These manifest our commitment to providing a high level of contributions to management accounting research and practice.

Marc J. Epstein
John Y. Lee
Editors

This page intentionally left blank

NON-FINANCIAL PERFORMANCE MEASURES IN THE HEALTHCARE INDUSTRY: DO QUALITY-BASED INCENTIVES MATTER?☆

John H. Evans, III, Andrew Leone and Nandu J. Nagarajan

ABSTRACT

This study examines the economic consequences of non-financial measures of performance in contracts between health maintenance organizations (HMOs) and primary care physicians (PCPs). HMOs have expanded contractual arrangements to give physicians not only financial incentives to control costs, but also to make the physicians accountable for the quality of patient care. Specifically, we examine how quality provisions in HMO–PCP contracts affect utilization (patient length of stay in the hospital), patient satisfaction, and HMO costs. Our results show that quality clauses are associated with a statistically significant increase in utilization (29 more hospital days annually per 1,000 HMO enrollees). Further, inclusion of quality clauses in PCP contracts also led to a significant increase in patient satisfaction, but no associated increase in

☆The data used in this study are available from the authors by request.

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 1–31

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14001-5

HMO costs. Overall, these results suggest that quality clauses in PCP contracts can increase value by increasing customer satisfaction without significantly increasing cost.

1. INTRODUCTION

Organizations use control systems to ensure that the agents they hire are accountable for their actions. Control systems typically provide agents with financial incentives based on specific financial, and more recently non-financial, performance measures tied to organizational goals. The choice of performance measures will depend on the nature of the incentive conflicts within the organization, the industry, and the competitive environment. Healthcare organizations face particular challenges in providing effective incentives because of regulated reimbursement and a complex value chain of relationships among hospitals, managed care organizations, employers, physicians, patients, and insurance companies.[1]

One prominent example of the development of a mix of performance measures arises when health maintenance organizations (HMOs) contract with primary care physicians (PCPs), specialists, and hospitals. The growth of the managed care industry is based primarily on their ability to provide employers with healthcare coverage for their employees at a lower cost. Many HMOs incorporate financial cost control incentives in their contracts with physicians, but as the managed care industry has grown, HMOs have experienced increasing political costs from the public's perception that emphasis on cost containment has reduced the quality of patient care. In response, a number of HMOs, beginning with U.S. Healthcare in 1987, have introduced quality-based financial incentives into their PCP contracts (Traska, 1988). The HMOs' objective in adding quality measures to physician contracts is to reinforce physicians' accountability for quality as well as cost.[2]

This trend in the healthcare industry toward combining non-financial and financial performance measures in physicians' incentive contracts follows similar practices elsewhere. Although publicly available data are limited in most industries, a few studies have analyzed the economic consequences of such performance-based incentives (e.g., Banker, Potter, & Srinivasan, 2000; Ittner, Larcker, & Rajan, 1997). An advantage of studying the healthcare industry is that HMO-reporting mandates in some states provide a data source on managed care physicians' contracts. We use these data on HMO–PCP contracts to produce results that address the sharp public

debate on managed care's use of incentives to influence physicians' trade-offs between the cost and quality of patient care.[3]

This study examines two potential economic consequences of including quality-based incentives in HMO–PCP incentive contracts. First, quality incentives may induce physicians to provide higher quality care to HMO patients. Second, the quality incentives may hinder other HMO cost control efforts. Consistent with both of these potential effects, PCPs indicate in survey responses that the inclusion of quality incentives in their contracts with HMOs significantly affects how they provide care to patients.[4] To test these perceptual findings, we provide archival evidence on these two potential economic consequences, and interpret the resulting trade-off between the effect of quality incentives on the cost and quality of patient care.

Prior studies, including Hillman, Pauly, and Kerstein (1989) Debrock and Arnould (1992), and Kerstein and Paik (1994), have focused on the fee-for-service (FFS) versus capitation feature of HMO–PCP contracts. Under an FFS arrangement, PCPs are compensated for services provided to enrollees based on an agreed-upon fee schedule. In contrast, under capitation for primary care services, PCPs are paid a fixed monthly fee, adjusted for age and sex for each enrollee, and independent of the amount of treatment provided by the PCP to the enrollee.[5] The prior studies find that capitation generally reduces resource consumption as proxied by the number of days HMO enrollees spend in the hospital.

Our study extends this previous work in several important ways. First, we update the sample period from approximately 1986 in prior studies to 1993, thereby capturing much of the recent growth in HMO enrollment.[6] Second, in addition to replicating the earlier studies using hospital days as a proxy for resource consumption, we extend the analysis of capitation versus FFS using more direct HMO cost measures. Third, we provide empirical evidence on the economic consequences of adding quality-based financial incentives to PCP contracts. This evidence demonstrates how quality incentives affect patient satisfaction, a relation not examined in earlier studies because the quality-based contractual features did not exist at that time. Finally, we also provide evidence that the HMOs achieved quality improvements without significantly increasing the cost of care, even though the quality incentives were associated with longer hospital stays. We discuss potential explanations for these seemingly inconsistent effects of quality provisions on the level of resource consumption.

This study is organized as follows. Section 2 provides additional background on HMO cost and quality incentive arrangements, and develops our hypotheses based on those incentives. Section 3 describes our sample data

and the method of testing our hypotheses. Section 4 reports empirical results and Section 5 provides our conclusions.

2. BACKGROUND AND HYPOTHESES DEVELOPMENT

Contracts between HMOs and PCPs influence the total cost of patient care because PCPs not only provide care to patients but also influence the cost of care provided by specialists and hospitals. PCPs can help to control the cost of specialists and hospitals by performing a "gatekeeper" function for HMOs. This involves working with hospitals to control both hospital admission and length of stay, which are important determinants of HMO healthcare costs. PCPs can influence the cost of care using such strategies as reducing hospital stays by teaching patients with congestive heart failure, diabetes, asthma, and other chronic diseases, how to manage these diseases themselves, reducing emergency room charges by providing extended PCP office hours, and reducing specialist referrals by providing certain types of more specialized care themselves.

Reducing the rate of hospital admissions should reduce the cost to HMOs of hospital services to their patients. Besides controlling hospital admissions, HMOs also seek to control hospital length of stay, particularly if the HMO's payment to the hospital includes a per diem charge for the hospital stay. Length of stay is an attractive performance measure because it is objectively measurable at a low cost, highly visible and well understood, and often responsive to a variety of administrative and clinical policy choices that physicians can control without materially altering the quality of patient care.[7] These relations have led HMOs to expend significant resources in the design and implementation of control systems to provide PCPs with incentive to control hospital admissions and hospital length of stay.[8] Although several features have contributed to the popularity of length of stay as a performance measure for physicians, cost reduction may ultimately depend on how effectively HMOs control the number and types of procedures that physicians perform (Evans, Hwang, & Nagarajan, 1995, 2001).

2.1. Financial Measures in PCP Contracts

HMOs employ both financial and non-financial performance measures in their control systems for PCPs. The two primary financial dimensions of

HMO–PCP contracts are whether the HMO pays the physician on an FFS versus a capitated basis and whether the HMOs include financial bonuses and penalties based on the extent of resource consumption in PCP contracts. Such contractual arrangements are particularly important because physicians are typically not employees of the HMOs to whom they provide services (Leone, 2002).[9] Instead, physicians operate as independent contractors, and therefore contractual incentives are likely to be even more important than in more common organizational environments in which other incentives such as promotion and termination can discipline employees.

We next discuss how an HMO's choices between contracting on an FFS versus a capitation basis with PCPs and whether or not to include bonuses for cost control in PCP contracts are likely to affect the cost of patient care. In the absence of other financial incentives, an FFS arrangement provides little or no incentive to the PCP to control the cost of services, tests, referrals, and hospitalization. In fact, PCPs frequently have incentives to increase hospitalization because this results in greater profit for the physician. Pauly (1980) argues that by hospitalizing a patient for a given treatment, the physician may be able to charge a higher fee for inpatient procedures, while simultaneously reducing the physician's cost of care because of support provided by the hospital.

Because capitated PCPs will realize profit equal to the difference between the fixed monthly capitation payments and their costs, they have an incentive to control the cost of care. PCPs can reduce their costs without reducing their revenue by reducing services to enrollees, including hospital days, so long as this reduction does not reduce patient satisfaction or increase the PCPs malpractice litigation cost.

Using data from a survey by Hillman (1987), Kerstein and Paik (1994) find that hospital utilization is greater when HMOs pay PCPs based on an FFS structure than when the HMOs make capitated payments to PCPs. This evidence is also consistent with the findings of Hillman et al. (1989) and Debrock and Arnould (1992). Similarly, Josephson and Karcz (1997) find significantly fewer hospital admissions for five specific ambulatory sensitive conditions under capitation than under traditional indemnity FFS. Although the proportion of HMOs using capitation has increased since 1986, we hypothesize that the incentive effect of capitation on the level of hospital utilization will continue to hold for our 1993 sample. Hence, Hypothesis H1 provides a replication of prior research, thereby establishing that our subsequent tests of the effect of quality clauses begin from common ground with the previous literature.

H1. Utilization of hospital resources will be smaller for HMOs using capitation contracts with their PCPs than for HMOs using FFS contracts.

Before moving to other financial incentives in PCP contracts, we first note two features of patient care that could lead to a lack of support for H1. First, capitation of PCPs could also potentially lead to overutilization of specialists' services. This unintended consequence would reflect PCPs reducing their own direct costs by referring patients to specialists rather than treating the patients themselves.[10] In turn, specialists who are paid on an FFS basis may then provide excessive services and hospitalization. To the extent that this effect is important in our sample, it will operate against finding empirical results consistent with H1.

Second, prior studies of the effect of capitation on resource utilization have generally not controlled for the proportion of Medicaid and Medicare enrollees in each HMO. Phelps (1992) reports evidence indicating that Medicare and Medicaid patients experience higher hospital utilization. Because HMOs with greater Medicare enrollment tend to make greater use of capitation contracts, the FFS versus capitation variable in prior studies may be reflecting the joint effect of the type of enrollees (proportion of Medicare and Medicaid) as well as the basic FFS versus capitation payment structure. To isolate the effect of FFS versus capitation, we examine the utilization of resources for only the non-Medicare/Medicaid patients of our sample HMOs.[11]

Besides capitation, a second financial performance measure used by some HMOs involves payment of a bonus for controlling HMO costs.[12] While capitation and cost control bonuses provide incentives for PCPs to control resource consumption, many HMOs attempting to retain and expand their market share have experienced competitive pressure to enhance the quality of their services.[13] In response, a number of HMOs have introduced new quality-based incentives into their contracts with PCPs, as we describe next.

2.2. Non-Financial Measures in HMO–PCP Control Systems

HMOs began to include quality provisions in PCP contracts following expressions of concern that contractual cost-control provisions reduced the quality of patient care.[14] The quality-based incentives generally pay PCPs bonuses if they score high enough on specified quality measures. Clearly, measuring the overall quality of patient care is very difficult (Blumenthal, 1996), and patient satisfaction scores will reflect many factors in addition to

the underlying appropriateness and technical expertise of care that the physician provides. For example, patient retention may be very sensitive to such features as the convenience of scheduling appointments and the courtesy of the physician's office staff, dimensions of care distinct from the technical quality of medical care provided.

In addition to measurement difficulties, another consideration in using quality bonuses is the potential for such incentives to increase HMO costs, including increased utilization of hospital resources. For example, PCPs with quality-based incentives may use more effective and more expensive drugs and treatment protocols, including hospitalization. Likewise, a PCP might endorse more hospitalizations and longer hospital stays to improve patient satisfaction, even if these changes were not very likely to improve the patient's ultimate medical condition. When HMOs use quality-based incentives to induce PCPs to provide higher quality care, the PCPs who are paid on a capitation basis but also receive quality-based bonuses face an explicit financial trade-off between cost and quality. That is, in addition to the general cost–quality trade-off faced by any service provider, a physician paid via both capitation and a quality bonus is literally being paid both to improve quality and to reduce cost.[15] Hence, relative to PCPs without quality-based incentives, physicians with quality bonuses have an additional rationale for increasing quality. If increasing the quality of care is positively associated with an increase in the level of service provided (Kerstein & Paik, 1994), PCPs with quality-based incentives may reduce utilization less than other PCPs. Further, to the extent that PCPs being paid quality-based incentives receive monetary returns from providing additional services, the HMOs paying such incentives should experience higher hospital utilization, as reflected in Hypothesis H2.[16]

H2. Utilization of hospital resources will be higher for HMOs that pay PCPs quality-based incentives than for HMOs that do not.

HMOs measure the overall quality of PCP patient care using member satisfaction surveys (MMSs), the frequency of delivery of specified preventive medicine procedures (e.g., mammograms, well-child checkups, etc.), patient transfer rates (the rate at which enrolled patients move from one PCP within the HMO to another PCP within the same HMO), medical chart reviews, and evaluations by the HMO administrative staff. The preventive care indices are referred to as HEDIS measures based on the set of standardized measures developed by the Health Plan Employer Data and Information Service (HEDIS). Gold, Hurley, Lake, Ensor, and Berenson (1995)

document that in 1994 over one-half of the HMOs in their national sample included quality-based incentives in their contracts with PCPs.

Patient satisfaction provides one overall measure of the quality of care that is likely to capture at least a portion of each of the dimensions described above. For this reason and because of data availability, we measure quality of care based on patient satisfaction scores. HMOs can then hold the PCPs responsible for the overall provision of services by including satisfaction measures in the PCPs' incentive compensation. To the extent that the PCP responds to these provisions, the quality-based incentives in the HMO–PCP contracts should be associated with improved patient satisfaction, as reflected in Hypothesis H3.

H3. Quality of patient care, as proxied by patient satisfaction measures, will be higher for HMOs that pay PCPs quality-based incentives than for HMOs that do not.

Hypothesis H2 focuses on hospital days as a prominent proxy for resource utilization. However, even if quality-based incentives lead to additional days of hospital care, these additional days of hospital care may not significantly increase HMO costs. Whether they do depends on the importance of patient days, as a volume-based cost driver, relative to other determinants of hospital and HMO costs such as the complexity of tests and procedures performed in treating patients. In another service industry context, Banker and Johnston (1993) find that, in addition to the volume of services, the complexity of services also significantly influences airline costs. In contrast, Foster and Gupta (1990) find no association between measures of complexity and overhead costs for their sample of manufacturing firms. In a healthcare context, Evans et al. (2001) provide evidence that the number and nature of procedures performed are the primary drivers of hospital costs, rather than the volume of patient days in the hospital. In deciding how to treat patients, physicians can control both patient length of stay and also the number and complexity of procedures used. To the extent that physicians seek to modify their practice patterns to earn a quality bonus without simultaneously triggering a significant increase in hospital costs, the preceding study's results suggest that modifying lengths of stay will be more effective than increasing the number and sophistication of procedures performed. This reasoning leads to the following hypothesis.

H4. Costs will be no greater for HMOs that pay PCPs quality-based incentives than for HMOs that do not.

However, we acknowledge that any increase in hospital days is likely to increase certain categories of hospital costs. To the extent that this is true, H4 is less likely to be supported. On the other hand, we also note that because H4 is stated in the null form, empirical results consistent with H4 could also be due to limitations on the power of our tests.

3. DATA, MODEL, AND VARIABLES

3.1. Data Sources and Description

We collected the data used in this study from three sources.

3.1.1. Data from State HMO Filings

First, our basic financial and contractual data, as well as information on HMO characteristics, come from HMO reports to state regulators. We collected HMO contract and financial performance data from the 1993 annual filings by HMOs to state regulators in eight states, selected on the basis of data availability and the cost of data collection. The filings generally contain financial, utilization, enrollment, and contracting information.

Specifically, we collected contractual data on the extent to which each HMO paid physicians on a capitated versus an FFS basis, and also whether the HMO paid bonuses to PCPs on the basis of cost control and quality measures. For our sample of HMO–PCP contracts, the most frequently used quality-based incentives are survey responses and medical chart reviews. Appendix A provides an example of an HMO–PCP contract with quality-based incentives. The example illustrates the potential economic importance of such quality incentives because the quality-based bonuses can increase the HMOs capitation payments to the physician by as much as 56%.

For the eight states, in which we collected HMO filing data (California, Illinois, Michigan, Minnesota, Ohio, Pennsylvania, Rhode Island, and Wisconsin), our data include 70 of 171 HMOs operating in those states in 1993. We excluded HMOs for which insufficient data were available from the state agency. For example, we have data for only three of 38 HMOs in California because most California HMOs chose not to make data available after they were exempted from the state's Freedom of Information Act.

Table 1 compares the 70 sample HMOs with the general population of 540 HMOs operating in the U.S. in 1993. The sample is generally similar to the population with respect to model type, tax status, size, and age. The

Table 1. Comparison of HMO Population Versus Sample.

	HMO Population				Sample			
Characteristics	Total members	% of total	Total HMOs	% of total	Total members	% of Total	Total HMOs	% of Total
All HMOs	39,783,140	100	540	100	9,038,486	100	70	100
Model type								
PGP	24,422,791	61.39	220	40.74	6,387,153	70.67 + f4/f2	32	45.71
IPA	15,360,349	38.61	320	59.26	2,651,333	29.33	38	54.29
Tax status								
NFP	19,382,127	48.72	176	32.59	3,500,384	38.73	26	37.14
For-profit	20,401,013	51.28	354	65.56	5,538,103	61.27	44	62.86
Size of HMO								
$\leqslant$ 4,999	122,848	0.31	60	11.11	2,191	0.02	2	2.86
5,000–14,999	848,971	2.13	88	16.30	100,019	1.11	10	14.29
15,000–24,999	1,622,027	4.08	82	15.19	183,045	2.03	8	11.43
25,000–49,999	4,611,431	11.59	130	24.07	611,987	6.77	19	27.14
50,000–99,999	5,759,851	14.48	81	15.00	968,853	10.72	13	18.57
$\geqslant$ 100,000	26,818,012	67.41	99	18.33	7,172,392	79.35	18	25.71
Age of HMO								
$<$ 1	79,158	0.20	15	2.78	0	0.00	0	0.00
1–2 years	256,057	0.64	19	3.52	0	0.00	0	0.00
3–5 years	1,542,348	3.88	57	10.56	31,623	0.35	2	2.86
6–9 years	11,400,002	28.66	257	47.59	5,010,056	55.43	36	51.43
$\geqslant$ 10 years	26,505,575	66.63	192	35.56	3,996,808	44.22	32	45.71
Region								
Northeast	9,462,961	23.79	100	18.52	2,664,598	29.48	11	15.71
South	7,706,017	19.37	169	31.30	0	0.00	0	0.00
Midwest	8,085,183	20.32	161	29.81	3,559,672	39.38	56	80.00
West	14,450,995	36.32	107	19.81	2,814,216	31.14	3	4.29

Note: PGP = Prepaid group practice, including all non-IPA HMO model types (network, mixed), IPA = Independent Practice Association.
Source: InterStudy Competitive Edge (1993).

most significant difference is in geographic representation, where the sample is highly concentrated in the Midwest (80% compared to 30% in the population). Accordingly, our subsequent analysis controls for the state in which the HMO operates.

3.1.2. Data from Other Sources

Our second source of data was the Area Resource File (ARF) from which we obtained market and control variables, as described in Appendix B. Third, we obtained survey data on patient satisfaction from "The NCQA Annual Member Health Care Survey", as reported in the National Research Corporation's (NRC) "NRC Report Card System", for health plan members for the period 1993–1994. The survey item asked HMO members, "How do you rate your overall satisfaction with your PCP?", using a five-point scale

(1 = lowest rating; 5 = highest rating). We obtained 2,941 survey responses from 37 of the HMOs in our sample. After dropping one of these HMOs due to missing data, our final sample includes 2,907 survey responses from 36 HMOs, with a minimum of 24 and a maximum of 439 responses per HMO.

3.1.3. Distribution of Contract Types

Table 2 shows the distribution of HMO–PCP contract types for the sample of HMO state filings used to test Hypotheses H1, H2, and H4. In addition to capitation versus FFS structure and the presence or absence of a quality bonus, Table 2 also displays the distribution of HMOs' tax status (for-profit versus not-for-profit (NFP)) and whether the contracts pay PCPs an additional bonus if utilization measures are within a budgeted target (utilization bonus). The 28 contracts represented in the first four rows of Table 2 all involve FFS payments to PCPs, while the 42 contracts in the last four rows employ capitation. Note that 39 of the 42 HMOs using capitation in our sample also employ utilization bonuses, suggesting that reducing utilization is an important consideration in the decision to use capitation. The small number of observations of HMOs employing capitation without utilization bonuses prevents us from distinguishing empirically between the effect of capitation alone versus capitation and utilization bonuses.[17] Therefore, our primary empirical analysis does not distinguish whether capitation contracts do or do not include utilization incentives.

Table 2. Frequency of HMO–PCP Contract Types in Sample ($n = 70$).

Contract Type	FP	NFP	Total
FFS			
No utilization bonus, no quality bonus	5	5	10
No utilization bonus, quality bonus	0	0	0
Utilization bonus, no quality bonus	11	6	17
Utilization bonus, quality bonus	1	0	1
CAP			
No utilization bonus, no quality bonus	3	0	3
No utilization bonus, quality bonus	0	0	0
Utilization bonus, no quality bonus	16	11	27
Utilization bonus, quality bonus	8	4	12
Total	44	26	70

Note: FP = for-profit HMO, NFP = not-for-profit HMO, FFS = fee-for-service, CAP = capitation.

Table 2 shows that 12 of the 42 sample HMOs using capitation combined it with a quality bonus, while only one HMO combined an FFS contract with a quality bonus. Further, quality-based bonus provisions are found in capitated and FFS contracts only in the presence of utilization-based incentive clauses, suggesting that HMOs may have included quality provisions in PCP contracts to avoid an excessive emphasis on cost control.

3.2. Model for Testing Hypotheses H1 and H2

We use the model in Fig. 1 to test Hypotheses H1 and H2 concerning HMO–PCP payment arrangements and HMO utilization of resources. The dependent variable, DAYS, the measure of utilization, is the annual number of days of inpatient hospital care per 1,000 group enrollees for that HMO.

3.3. Variable Descriptions for the Hospital Days Regression

The model of hospital utilization in Fig. 1 captures two dimensions of HMO–PCP contracts. First, CAPITATION is 1 if PCPs are paid on a capitated basis, and 0 otherwise. They may or may not receive utilization bonuses. Second, QUALITY is 1 if PCPs are eligible to receive a quality-based bonus,

$$\text{DAYS} = \beta_0+\beta_1\text{CAPITATION}+\beta_2\text{QUALITY}+\beta_3\text{AVPREM}+\beta_4\text{IPA}+\beta_5\text{DOCCAP}+\beta_6\text{MKTPWR}+\beta_7\text{BEDSi}+\beta_8\text{NFP}+\beta_9\text{PERMED}+\varepsilon$$

Dependent variable: **DAYS = Hospital Utilization per 1000 group enrollees**	
Independent Variables	Predicted Sign
CAPITATION = 1 if capitation; 0 otherwise	-
QUALITY = 1 if quality incentive; 0 otherwise	+
AVPREM = Average HMO monthly premium	+
IPA=1 If HMO is IPA or Network; 0 otherwise	+
DOCCAP= Number of PCPs per capita	-
MKTPWR = Market penetration by HMOs	-
BEDS = Number of hospital beds per capita	+
NFP = 1 if not-for-profit HMO; 0 otherwise	+
PERMED = Log of the proportion of Medicare and Medicaid enrollment in the HMO	?

Fig. 1. Determinants of HMO Members' Utilization of Hospital Services.

and 0 otherwise. Again, they may or may not receive a utilization bonus. We coded an HMO–PCP contract as including a quality bonus only if the contract explicitly described the bonus arrangement.

In addition to the preceding hypothesized variables, prior research identifies the following control variables as potentially associated with utilization. AVPREM is the estimated average per capita premium received by each HMO during 1993. We estimate the premium by dividing the HMO's total 1993 premium revenue by the total number of member months for 1993. HMOs may command higher premiums by providing more extensive patient care, which would then be reflected in higher rates of utilization. Consequently, HMOs with higher premiums are expected to have higher utilization, and we include AVPREM to control for this potential alternative explanation for variation in the number of hospital days. Next, previous research (Welch, 1988; Miller & Luft, 1995) find that, on average, HMOs organized as Independent Practice Associations (IPAs) experience higher utilization rates than do other HMO model types. Miller and Luft attribute the weaker controls over resource consumption in IPAs to the following. First, IPAs invest less in identifying physicians with conservative treatment styles than do staff and prepaid group practice (PGP) HMOs. Second, group norms are less effective at controlling resource consumption in an IPA environment where physicians typically practice independently. Finally, revenue from HMOs (at the time of our sample) is likely to account for a much smaller percentage of the physician's total revenue in an IPA environment, thereby weakening the PCPs cost control incentives. Therefore, the dummy variable IPA is included to control for this hypothesized effect of the HMO's organizational form.

We construct the market variables DOCCAP, MKTPWR, and BEDS from the ARF using weighted averages of county HMO enrollments. DOCCAP reflects the supply of physicians, and as this variable increases, physicians' bargaining power relative to HMOs is likely to decline, enabling HMOs to specify the contract form that they prefer, and thereby resulting in lower utilization. MKTPWR measures an HMO's market power, and as this power increases, the HMO is again more likely to be able to specify HMO–PCP contractual terms favorable to the HMO, consistent with reduced utilization. In markets with larger values of BEDS, hospitals facing greater financial pressure to utilize excess capacity may respond by increasing the average length of stay in the hospital by HMO enrollees. Consequently, we include BEDS to control for this potential supply-side effect on utilization.

Next, NFP is coded 1 if the HMO's tax status is NFP and 0 otherwise. PERMED is a log transformation of the proportion of Medicare and

Medicaid enrollees served by the HMO. Although our analysis focuses on the effect of PCP contracts on care provided to employee group enrollees (not Medicare or Medicaid enrollees), HMO utilization for group enrollees could still potentially be influenced by the extent of the HMO's Medicare and Medicaid enrollment, as discussed in note 9. Medicaid enrollment data are from annual HMO filings to state agencies, and Medicare data are from the Health Care Financing Administration's (HCFA) Medicare Report.

Table 3 (Panel A) provides descriptive statistics for the variables included in the sample of HMO state filings. The median number of hospital days per 1,000 enrollees is 295 and the median monthly HMO premium is $123. Mean values indicate that 60% (42) of the sample HMOs use capitation in PCP contracts, 18.6% (13) use quality-based financial incentives for PCPs, 54.3% (38) HMOs are organized as IPAs, and 37.1% (26) are NFP. Table 3 (Panel B) provides a correlation table.

4. EMPIRICAL RESULTS

4.1. Hypothesis H1

Using the regression results in column (a) of Table 4, we test H1 based on the estimated coefficient of –37.685 on CAPITATION in column (a), which is significantly different from zero at the $p<0.01$ level (one-tail test).[17] This coefficient indicates that, compared to FFS, using capitation for PCPs is associated with an average reduction of 38 hospital days per 1,000 enrollees per year, a decline of 12% compared to FFS. This result confirms the hypothesized effect of capitation from prior research and lays the groundwork for this paper's analysis of the incremental effect of quality-based incentives on the number of hospital days consumed.

Results in Table 4 are provided both with (columns a and b) and without (column c) control variables for the state in which the HMO operates and for national HMO firms, i.e., firms that operate HMOs or "plans" in multiple states. We control for state to capture any potential variation in practice patterns across geographic regions that may lead to differences in utilization (Wennberg, 1984; Diehr et al., 1990; Schwartz et al., 1994). We also control for national HMO firms because multiple observations from the same national HMO may not be independent. In our study, an HMO observation represents an HMO (plan) registered to do business in a state. A national HMO firm may own a number of HMOs (plans) throughout the country. Aetna, Metlife, Prudential, and TakeCare are the four national

Table 3. Descriptive Statistics for the Sample of HMO State Filings ($n = 70$).

Variable	Mean	Median	S.D.	Min	Max
Panel A					
DAYS	302.412	295.437	52.511	210.299	456.384
CAPITATION	0.60	1	0.493	0	1
QUALITY	0.186	0	0.391	0	1
AVPREM	125.089	123.236	24.813	67.194	239.677
IPA	0.543	1	0.5018	0	1
DOCCAP	0.029	0.0285	0.010	0.007	0.0553
MKTPWR	0.069	0.0395	0.114	0.012	0.904
BEDS	0.004	0.004	0.001	0.002	0.007
NFP	0.371	0	0.487	0	1
PERMED	−6.59	−4.233	4.883	−11.513	−0.102

Panel B – Pearson Correlations

Variable	Capitation	Quality	AVPREM	IPA	DOCCAP	MKTPWR	BEDS	NFP	PERMED
CAPITATION	1	0.37***	−0.10	−0.15	−0.03	0.09	0.06	−0.05	−0.27
QUALITY		1	−0.45	−0.03	0.02	0.27**	0.18	−0.04	−0.10
AVPREM			1	−0.00	−0.04	−0.19	0.07	−0.04	0.17
IPA				1	−0.09	−0.15	0.00	−0.12	−0.06
DOCCAP					1	0.05	0.24**	0.21*	0.03
MKTPWR						1	−0.01	−0.05	0.10
BEDS							1	−0.03	0.13
NFP								1	0.15
PERMED									1

Note: DAYS is the number of hospital days per 1,000 HMO enrollees in the fiscal year. CAPITATION is 1 if PCPs are paid on a capitated basis, and 0 otherwise. QUALITY is 1 if PCPs are eligible to receive a quality-based bonus, and 0 otherwise. AVPREM is the estimated average per capita premium received by each HMO during 1993. IPA is 1 if the HMO is organized as an IPA and 0 otherwise. DAYS, CAPITATION, AVPREM, and IPA are obtained from HMO reports to state agencies. DOCCAP, MKTPWR, and BEDS are obtained from the ARF, using weighted averages of county HMO enrollments. DOCCAP is the number of doctors per capita in the HMO's market area. MKTPWR is the percentage of the local area population enrolled in HMOs divided by the number of HMOs in the area. BEDS is the number of general hospital beds per capita. NFP is coded 1 if the HMO's tax status is NFP, and 0 otherwise. PERMED is a log transformation of the proportion of Medicare and Medicaid enrollees served by the HMO.

*Significant at 0.10 level of significance.

**Significant at 0.05 level of significance.

***Significant at 0.01 level of significance.

Table 4. Determinants of HMO Utilization of Hospital Resources (Dependent Variable is Hospital Days per 1,000 Group Enrollees).

Independent Variables	(a) $n = 70$ Coefficient (t-Statistic)	(b) $n = 70$ Coefficient (t-Statistic)	(c) $n = 70$ Coefficient (t-Statistic)
INTERCEPT	313.896 (7.003)***	274.011 (6.462)***	261.324 (7.385)***
CAPITATION (1 if capitation; 0 if FFS)	−37.685 (3.798)***	−36.897 (3.653)***	−29.906 (2.503)***
QUALITY (1 if quality-based incentive; 0 otherwise)	28.721 (1.947)**	22.514 (1.672)**	18.991 (1.481)
AVPREM (average premium)	−0.021 (0.077)	0.258 (1.053)	0.180 (0.948)
IPA (1 if HMO is an IPA model type; 0 otherwise)	−13.682 (1.306)	−11.074 (1.127)	−1.673 (0.161)
DOCCAP (the number of doctors per capita in the HMO service area)	−526.075 (0.896)	−761.168 (1.159)	−978.685 (1.775)*
MKTPWR (a measure of the market power of the HMOs in the HMO service area)	−66.553 (1.348)	−41.562 (1.310)	−29.845 (0.918)
BEDS (the number of hospital beds per capita in the HMO service area)	11,985 (2.849)***	13,246 (2.907)***	15,179 (2.953)***
NFP (1 if the HMO is a not-for-profit HMO; 0 otherwise)	59.310 (4.793)***	59.301 (5.281)***	40.628 (3.545)***
PERMED (log of the proportion of Medicare and Medicaid enrollment)	3.067 (2.881)***	2.290 (2.369)**	1.831 (1.742)*
(State dummies)			
IL	7.506 (0.465)	(0.394)	—
MI	−35.152 (1.663)*	−37.699 (1.836)*	—
N	−90.967 (3.573)***	−74.704 (2.638)***	—
WI	−18.005 (1.077)	−13.229 (0.801)	—
OH	−11.357 (0.607)	−8.349 (0.502)	—
CA	−41.930 (2.633)***	−42.429 (2.648)***	—

Table 4. (*Continued*)

Independent Variables	(a) $n = 70$ Coefficient (*t*-Statistic)	(b) $n = 70$ Coefficient (*t*-Statistic)	(c) $n = 70$ Coefficient (*t*-Statistic)
RI	−59.985 (2.733)***	−51.683 (2.379)**	—
(National firm dummies)			
Aetna	−3.537 (0.190)	—	—
Metlife Health Care Management	56.003 (2.198)**	—	—
Prudential Health Care Plans	8.591 (0.308)	—	—
TakeCare, Inc.	−9.537 (0.630)	—	—
Adjusted R^2 *F*-statistic	0.35 2.876	0.35 3.309	0.306 4.383

Note: The level of analysis is the HMO plan. IL, MI, MN, WI, OH, CA, RI are state dummy variables to control for variation arising from local practice patterns and state regulation. Aetna, Metlife, Prudential, and TakeCare, Inc. are National HMO dummies. Refer to Table 3 for details of variable construction and data sources. One-tailed tests are used for the hypothesized explanatory variables, CAPITATION and QUALITY, and two-tailed tests are used for all other variables (White corrected standard errors).
(a) – This model controls for both the state in which the HMO operates and for national HMOs that own multiple plans.
(b) – This model controls for the state but not for national HMOs.
(c) – This model does not control for state or national HMOs.
*Significant at 0.10 level of significance.
**Significant at 0.05 level of significance.
***Significant at 0.01 level of significance.

HMOs in our sample. Previous studies on HMO–PCP contracts have treated each HMO (plan) as a separate independent observation (Hillman et al., 1989; Kerstein & Paik, 1994; Wholey, Feldman, & Christianson, 1995), creating a potential omitted variables problem and inflating significance test statistics if observations are not independent.

The results in columns (b) and (c) of Table 4 demonstrate that the results from column (a) are robust to dropping the state and national HMO dummy variables, although the results are strongest when both controls are included.

Other variables that are positively associated with utilization for the model specification in column (a) of Table 4 are the number of BEDS,

not-for-profit HMOs (NFP), and the proportion of Medicare and Medicaid enrollment (PERMED). These results are consistent with greater utilization in markets with excess hospital bed capacity, in NFP HMOs, and in HMOs with a greater proportion of Medicare/Medicaid enrollees. Coefficients on the state dummy variables for Michigan, Minnesota, California, and Rhode Island are all negative and statistically significant, consistent with lower utilization rates in states such as Minnesota and California, which have pioneered the development of managed care.

4.2. Hypothesis H2

H2 hypothesizes that including quality-based incentives will result in increased resource utilization as measured by the number of patient days of hospital care. We test H2 based on the estimated coefficient of QUALITY in column (a) of Table 4, which is $+28.7$, significant at the $p = 0.028$ level (one-tail test). This coefficient indicates that quality-based incentives are associated with an additional 28.7 hospital days per 1,000 enrollees per year, an increase of 9% over the overall mean number of hospital days per 1,000 group enrollees. These results indicate that quality-based incentives are associated with significantly greater utilization in both a statistical and an economic sense. In turn, compared to FFS, the estimated combined effect of capitation and quality becomes the sum of the estimated coefficients of CAPITATION and QUALITY, which is –9.0.[18] A χ^2 test fails to reject the null that the combined effect of CAPITATION and QUALITY is equal to zero.

4.3. Hypothesis H3

H3 hypothesizes that quality bonuses in PCP contracts will be associated with increased patient satisfaction relative to contracts without quality-based incentives. To test H3, we use the results in Table 5 for a regression of individual patient satisfaction responses on HMO–PCP contract characteristics, specifically CAPNQ (CAPNQ = 1 for HMOs using capitation but no quality incentives; 0 otherwise) and CAPANDQ (CAPANDQ = 1 for HMOs using capitation and quality incentives; 0 otherwise). Control variables include AVPREM, IPA (dummy variable for IPA HMO form), NFP (dummy variable for not-for-profit HMO), MKTPWR, DOCCAP, and BEDS.

Table 5. Patient Satisfaction as a Function of HMO–PCP Contract Features (Dependent Variable is Patient Satisfaction Score on a Five-Point Scale with 5 = Highest and 1 = Lowest).

Variable	Description	Model (*t*-Statistic)
INTERCEPT	Intercept	2.826 (1.431)
CAPNQ	Capitation and no quality incentive	0.072 (0.258)
CAPANDQ	Capitation and quality incentive	0.448** (2.185)
AVPREM	Average premium	0.014 (1.188)
IPA	1 if IPA and 0 otherwise	0.948*** (3.842)
NFP	1 if NFP and 0 otherwise	1.438*** (6.387)
MKTPWR	Market power	−16.790** (−2.181)
DOCCAP	Doctors per capita	4.435*** (3.034)
BEDS	Beds per capita	−61.303*** (−3.744)
	Fixed effects?	Yes
	Adjusted R^2	0.116
	Number of observations	2,907

Note: The level of analysis for this regression is the survey respondent. The dependent variable, customer satisfaction, is the survey respondent's response to the following question: "How do you rate your overall satisfaction with your PCP?", using a five-point scale (1 = lowest rating; 5 = highest rating). The regression includes HMO fixed effects. See Table 3 for more details on variable construction. One-tailed tests are used for the hypothesized explanatory variables, CAPNQ and CAPANDQ, and two-tailed tests are used for all other variables (White corrected standard errors).

*Significant at 0.10 level of significance.

**Significant at 0.05 level of significance.

***Significant at 0.01 level of significance.

The coefficient on CAPANDQ is +0.448, which is significantly different from 0 at the 0.05 level (one-tail). The estimated coefficient indicates that, compared to FFS contracts, capitation contracts with quality-based incentives increase patient satisfaction by almost one-half point, a 12% increase relative to the overall mean patient satisfaction survey response of 3.67 for our sample.

We ran a variety of robustness checks on the results in Table 5. For 12 alternative specifications constructed by dropping control variables from Table 5, and in some cases adding the control variable PROPOLD (proportion of HMO patients over 65), CAPANDQ is positive eleven times and is significant at the 0.05 level (one-tail) five times. The primary pattern in these results is that when we exclude (include) PROPOLD, CAPANDQ is (is not) statistically significant. Checking for the association between

CAPANDQ and PROPOLD reveals that HMOs using capitation and quality incentives (CAPANDQ = 1) have a significantly greater proportion of older patients than do other HMOs. Specifically, the proportion of enrollees over age 65 is 12.7% when CAPANDQ = 1 versus 11.0% when CAPANDQ = 0. A t-test rejects the null that these proportions are equal at the $p = 0.001$ level. This positive association between CAPANDQ and PROPOLD appears to explain why CAPANDQ declines in significance when we include PROPOLD in the regressions.

These results suggest that, subject to some modest concerns with robustness, we find reasonable support for the hypothesis that including quality incentives in HMO–PCP contracts leads to increased patient satisfaction.

4.4. Hypothesis H4

Finally, we test H4, which proposes that although including quality incentives in PCP contracts may lead to increased hospital days (H3), HMOs will not experience a corresponding increase in the cost of providing care. The results in Table 6 are for a regression of HMO costs, measured as the ratio of the HMO's medical expenses to the HMO's total revenue, on characteristics of HMO–PCP contracts (CAPNQ, CAPANDQ), other HMO characteristics (AVPREM, IPA, PERMED, NFP), and control variables reflecting features of the HMO's market, DOCCAP, BEDS, WINCAP (income per capita), and WINFECTIONCAP (the rate of infectious disease per capita).

Both capitation variables (CAPNQ, CAPANDQ) have coefficients that are negative and statistically significant, indicating that the capitation feature, independent of the quality bonuses, is associated with lower HMO cost than is the FFS feature. To assess the incremental effect of the quality bonuses, we note that the coefficient on CAPANDQ is a larger absolute value than that on CAPNQ, although the difference is not statistically significant ($F = 0.04$; $p > 0.83$). This result implies that adding quality incentives do not increase HMO costs for our sample, consistent with H4.

5. CONCLUSION

This paper examines the economic consequences of incorporating a nonfinancial performance measure in an HMOs' contracts with PCPs. In particular, we focus on the effect of using capitation versus FFS to compensate

Table 6. HMO Cost as a Function of HMO–PCP Contractual Features (Dependent Variable is the Ratio of HMO Medical Expense to HMO Total Revenue).

Variable	Description	(a) Model 1 (*t*-Statistic)	(b) Model 2 (*t*-Statistic)
INTERCEPT	Intercept	1.057***	1.115***
		(12.785)	(7.618)
CAPNQ	Capitation and no quality incentive	−0.060***	−0.059***
		(−2.731)	(−2.66)
CAPANDQ	Capitation and quality incentive	−0.064**	−0.066**
		(−2.094)	(−2.125)
AVPREM	Average premium	−0.002***	−0.002***
		(−4.000)	(−3.698)
IPA	1 if IPA; 0 otherwise	−0.061***	−0.060***
		(−2.863)	(−2.764)
PERMED	Percent Medicare enrollees	0.001	0.001
		(0.671)	(0.636)
NFP	1 if NFP; 0 otherwise	0.012	0.009
		(0.483)	(0.358)
MKTPWR	Market power	−0.059	−0.047
		(−0.647)	(−0.489)
DOCCAP	Doctors per capita	0.262	0.257*
		(1.943)	(1.878)
BEDS	Beds per capita	−0.429	−0.713
		(−0.432)	(−0.663)
WINCAP	Income per capita		0.000
			(−0.744)
WINFECTIONCAP	Infectious disease per capita		0.240
			(0.78)
	Adjusted R^2	0.28	0.26
	Number of observations	69	69

Note: The level of analysis is the HMO plan. Column (b) reports results after including two additional variables. WINCAP is income per capita in the HMO's market area. WINFECTIONCAP is the rate of infectious disease per capita reported in the HMO's market area. These variables are obtained form the ARF. See Table 3 for details of variable construction and data sources for the other variables. One-tailed tests are used for the hypothesized explanatory variables (CAPNQ and CAPANDQ), and two-tailed tests are used for all other variables (White corrected standard errors).

*Significant at 0.10 level of significance.

**Significant at 0.05 level of significance.

***Significant at 0.01 level of significance.

PCPs, and the inclusion of financial incentives based on quality measures in PCP contracts. Using data from HMO filings with state agencies, as well as patient satisfaction survey results, we test four hypotheses.

First, our results support H1, which replicates previous studies' finding that contracts that base payments to PCPs on a capitation basis result in lower resource consumption (an estimated 38 fewer hospital days per year per 1,000 enrollees) than under an FFS arrangement. Our second hypothesis, H2, proposes that HMOs that pay PCPs quality-based incentives will experience a higher level of resource consumption measured by hospital days than will HMOs that do not pay such bonuses. Regression results indicate that quality-based incentives are associated with an estimated increase of 29 days of hospital care per year per 1,000 enrollees, consistent with H2. This increase in hospital days offsets most, but not all, of the estimated effect of capitation in reducing hospital days.

Third, H3 proposes that contracts with quality-based financial incentives for PCPs will result in higher levels of patient satisfaction than contracts without quality-based incentives. Our results based on HMO patient surveys provide support for H3, although the results are not entirely robust to which control variables we include in the regressions. Finally, in H4 we tested whether the increase in hospital days associated with quality provisions in PCP contracts translated to a corresponding increase in HMO costs. Our results indicate that capitation reduces HMO costs significantly, but adding quality provisions to a contract with capitation produces no significant change in costs.

Taken together, our empirical results suggest that quality provisions do have significant economic effects. Quality provisions in PCP contracts are associated with increased patient perceptions of HMO service quality, consistent with results from surveys of PCPs indicating that the quality incentives do affect their medical practice patterns (see note 3). Besides producing this desirable effect for HMOs, quality bonuses are estimated to produce no significant increase in HMO costs even though total hospital days do increase. These results suggest that HMOs have successfully used quality provisions to offset some of the negative effects of capitation, while largely preserving the cost saving effects of capitation.

Several limitations of our study should be noted. First, in some cases, HMO–PCP financial contracting involves two levels of contracts, the first a contract between an intermediate entity such as a physician group, and a second contract between that entity and individual physicians (Welch, Hillman, & Pauly, 1990). Our analysis captures only the form of the HMO–physician group contract, and not the physician group-PCP contractual

form, when such an additional layer exists. We are unable to estimate the effect of omitting the HMO–physician contracts. A second limitation of our approach is that patient satisfaction results are available for only 36 of the 70 HMOs in our original sample. Finally, we use three separate OLS equations to estimate the determinants of utilization, quality, and cost. Although beyond the scope of the present study, a more general approach would these variables to be simultaneously determined in a system of equations.

NOTES

1. For details, see Herzlinger (1997).
2. HMOs' implementation of systems to hold physicians accountable for quality as well as cost illustrates Epstein and Birchard's (1999) argument for the advantages of institutionalizing accountability. The path that HMOs have followed parallels their recommendation to develop new performance measures, incorporate them into integrated management systems, and then report the results, although primarily internally rather than externally.
3. See, for example, "HMOs' Woes Reflect Conflicting Demands of American Public", WSJ, 12/22/97, p. 1, and "Patients or profits?", The Economist, 3/7/98, pp. 15, 23–26.
4. A national sample of PCPs rated how much their medical practice patterns were affected by the results of patient satisfaction surveys. PCPs with quality bonuses in their contracts rated the effect as significantly greater than did PCPs without quality bonuses. Similar differences exist for quality measures other than patient satisfaction survey results.
5. A 1992 InterStudy survey found that 29% of all HMOs reimbursed PCPs exclusively on a capitated basis. In addition, PCPs paid primarily on a capitated basis may also receive part of their pay on a FFS basis. A review of over 100 PCP agreements submitted by HMOs to various state agencies during our sample period indicates that approximately 75% of the HMOs that pay their PCPs primarily on a capitated basis also make some use of FFS payments. In such cases, PCPs are compensated for most services on a capitated basis, but certain tests and infrequent services remain on an FFS basis. We treat such cases as capitated contracts in the subsequent empirical analysis.
6. During the period from 1986 to 1993, total HMO enrollment grew from 7.9% to 15.3% of the U.S. population (InterStudy Competitive Edge, 2001 and private communication, 2002].
7. For example, many hospitals have implemented "critical pathways", which specify detailed operational steps for a procedure such as a knee transplant to eliminate patient waiting time, redundancy and unnecessary process variations. The detailed study and restructuring of operational procedures frequently reduces a typical patient's hospital stay by several days.
8. An extensive literature in health economics and health services research uses hospital days per 1,000 enrollees as a proxy for cost, for example, Kerstein and

Paik (1994), Debrock and Arnould (1992), Miller and Luft (1994, 1995) and Hillman et al. (1989). In his analysis of the HMO industry, Luft (1987) states that, "…hospital utilization is the primary focus of HMO attempts to control cost" (p. 76), and further that, "The best overall measure of hospital use is the number of inpatient days per 1,000 enrollees per year" (p. 85). In fact, the growth and development of HMOs in the last 25 years, and government support of that growth, has relied on controlling the growth of health care costs, in part, by relying on HMOs to reduce the number of hospital days consumed (Carey, Mazo, Meyerson, & Edelman, 1972). Further evidence of the importance of inpatient care as a determinant of hospital costs is provided by Wrightson (1990), who notes that in HMO-physician contracts, inpatient hospital expense is typically the single largest expense category.

9. The exception is staff model HMOs in which physicians are HMO employees who are typically paid on a salary basis. In most cases, we exclude all such staff model HMOs from our empirical analysis.

10. Some more recent HMO contracts have been constructed as full risk capitation where the physician receives a capitation payment for all healthcare services, including the use of specialists, and is responsible for such costs. However, this arrangement is more common in group practice HMOs, and such groups ultimately do not pass this entire risk on to the PCPs. HMOs have also experimented with "reverse capitation" arrangements in which specialists receive capitated payments while PCPs have FFS contracts (Jaklevic, 1997).

11. Although our statistical tests reflect only the non-Medicare/Medicaid patients, some of the HMOs in our sample also treat Medicare and Medicaid patients. Because the presence of such patients could potentially influence the cost of treating non-Medicare/Medicaid patients, we also control for the volume of Medicare/Medicaid patients that each HMO treats.

12. We describe later why our data does not permit us to test for the independent effect of cost control bonuses in PCP contracts on the resulting cost of care.

13. For example, see "Quality is Focus for Health Plans" (New York Times, July 3, 1995), which describes how employers perceive that "much has been done to lower costs but now it's time to look at quality".

14. The development and use of patient satisfaction measures follows the broader trend in U.S. business in which firms emphasize customer satisfaction as a strategic management imperative (Anderson & Mittal, 2000). According to this management strategy, an HMO's performance on various attributes determines patient (customer) satisfaction, which then drives patient retention, a key determinant of HMO profitability. A key tenet of this approach is that retaining a current patient is much less expensive than acquiring a new patient (Anderson & Mittal, 2000), making customer retention a key driver of firm profitability. In turn, because research has established that greater customer satisfaction has a significant positive effect on customer retention (Bolton, 1998), HMOs can identify and manage those features of HMO-patient interaction that determine patient satisfaction.

15. PCPs who can also qualify for an additional bonus (or penalty) based on cost control face an additional trade-off.

16. Besides seeking to increase the quality of care itself, HMOs were also motivated to introduce quality-based incentives to reduce associated political pressure stemming from concerns that HMOs focused exclusively on profit at the cost of

patient care. Evidence of the political pressure that HMOs faced included the provision in the Omnibus Budget Reconciliation Act (OBRA) of 1986 that would have prohibited HMOs from making incentive payments to physicians to reduce or limit services to enrollees, as well as other state-level HMO regulatory legislation enacted in at least 40 states. However, whether HMOs acted out of concern for the quality of care or out of concern with regulatory costs, in either case implementing the quality clauses creates an additional incentive for physicians to increase hospital utilization.

17. For all of our model specifications, we estimate Variance Inflation Factors (VIF) to test whether multicollinearity appears to be problematic. The VIF values range from 1 to 3 in all of our models, suggesting that multicollinearity is not a serious problem.

18. Ideally, in addition to the tests for the effects of capitation (H1) and quality incentives (H2), we would also test for the interaction of the two. That is, do quality incentives have a significant effect when added to capitation contracts versus when added to FFS contracts? However, the observed empirical association of capitation and the quality incentives in our sample is so strong as to preclude our using this form of a test. Of our 13 observations of HMOs with quality incentives, 12 also use capitation. This means that the interaction term for capitation and quality would be absolutely identical to the quality variable (QUALITY) for 69 of our 70 observations, and this high degree of colinearity prevents us from identifying both a main effect of quality and also a capitation-quality interaction effect.

19. MSA (Metropolitan Statistical Area) is a code assigned by the Bureau of the Census which identifies a metropolitan area. The MSA can include one or more counties for each metropolitan area.

ACKNOWLEDGMENTS

We thank two anonymous reviewers and the editor for helpful comments and suggestions. Financial support from the Center for Research on Contracts and the Structure of Enterprise, Katz Graduate School of Business, University of Pittsburgh and from the Richard D. Irwin Foundation is gratefully acknowledged.

REFERENCES

Anderson, E., & Mittal, V. (2000). Strengthening the satisfaction-profit chain. *Journal of Service Research*, *3*(2), 107–120.

Banker, R. D., & Johnston, H. H. (1993). An empirical study of cost drivers in the U.S. airline industry. *Accounting Review*, *68*(3), 576–601.

Banker, R. D., Potter, G., & Srinivasan, D. (2000). An empirical investigation of an incentive plan that includes nonfinancial performance measures. *Accounting Review*, *75*(1), 65–93.

Blumenthal, D. (1996). Part 1: Quality of care – what is it? *New England Journal of Medicine*, *335*(12), 891–895.

Bolton, R. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. *Marketing Science*, *17*(1), 45–65.

Carey, C., Mazo, J., Meyerson, B., & Edelman, T. (1972). *Health maintenance organizations: An introduction and survey of recent developments*. Washington, DC: Lawyers' Committee for Civil Rights Under Law.

Debrock, L., & Arnould, R. (1992). Utilization control in HMO's. *Quarterly Review of Economics and Finance*, *32*(3), 31–53.

Diehr, P., Cain, K., Connell, F., & Volinn, E. (1990). What is too much variation: The null hypothesis in small-area analysis. *Health Services Research*, *24*, 2741–2771.

Epstein, M. J., & Birchard, B. (1999). *Counting what counts: Turning corporate accountability into competitive advantage*. Reading, MA: Perseus Books.

Evans, J. H., Hwang, Y., & Nagarajan, N. J. (1995). Physicians' response to length -of- stay profiling. *Medical Care*, *33*(11), 1106–1119.

Evans, J. H., Hwang, Y., & Nagarajan, N. J. (2001). Management control and hospital cost reduction: Additional evidence. *Journal of Accounting and Public Policy*, *20*(1), 74–90.

Foster, G., & Gupta, M. (1990). Manufacturing overhead cost driver analysis. *Journal of Accounting and Economics*, *12*, 309–337.

Gold, M. R., Hurley, R., Lake, T., Ensor, T., & Berenson, R. (1995). A national survey of the arrangements managed-care plans make with physicians. *New England Journal of Medicine*, *333*, 1678–1683.

Herzlinger, R. E. (1997). *Market-driven healthcare: Who wins, who loses in the transformation of America's largest service industry*. Reading, MA: Addison-Wesley Publishing Co. Inc.

Hillman, A. L. (1987). Financial incentives for physicians in HMOs: Is there a conflict of interest? *The New England Journal of Medicine*, *317*, 1743–1748.

Hillman, A. L., Pauly, M. V., & Kerstein, J. J. (1989). How do financial incentives affect physicians' clinical decisions and the financial performance of health maintenance organizations? *The New England Journal of Medicine*, *321*, 86–92.

Ittner, C., Larcker, D., & Rajan, M. (1997). The choice of performance measures in annual bonus contracts. *Accounting Review*, *72*(2), 231–255.

Jaklevic, M. C. (1997). Reversing capitation. *Modern Healthcare*, June 9, 4–15.

Josephson, G. W., & Karcz, A. (1997). The impact of physician economic incentives on admission rates of patients with ambulatory sensitive conditions: An analysis comparing two managed care structures and indemnity insurance. *Journal of Managed Care*, *3*(1), 49–56.

Kerstein, J., & Paik, T.Y. (1994). Does the method of compensation affect the firm's performance?: A study of payment methods in the prepaid health care industry. Presented at the NYU Accounting Seminar Series. (September). New York University: New York City.

Leone, A. (2002). The relation between efficient risk-sharing arrangements and firm characteristics: Evidence from the managed care industry. *Journal of Management Accounting Research*, *14*, 97–117.

Luft, H. S. (1987). *Health maintenance organizations: Dimensions of performance*. Somerset, NJ: Transaction Publishers.

Miller, R. H., & Luft, H. S. (1994). Managed care plan performance since 1980. *Journal of the American Medical Association*, *271*(19), 1512–1519.

Miller, R. H., & Luft, H. S. (1995). Estimating health expenditure growth under managed competition: Science, simulations, and scenarios. *Journal of the American Medical Association*, *273*(8), 656–662.
Pauly, M. (1980). *Doctors and their workshops*. Chicago: National Bureau of Economic Research, The University of Chicago Press.
Phelps, C. E. (1992). *Health economics*. New York, NY: HarperCollins Publishers, Inc.
Schwartz, M., Ash, A., & Anderson, J. (1994). Small area variations in hospitalization rates: How much depends on how you look. *Medical Care*, *32*, 1889–2201.
Traska, M. R. (1988). HMO uses quality measures to pay its physicians. *Hospitals*, *62*(13), 34–35.
Welch, W. P. (1988). IPAs emerge on HMO forefront. *Business and Health*, *5*(4), 14–16.
Welch, W. P., Hillman, A. L., & Pauly, M. V. (1990). Toward new typologies for HMO's. *Milbank Quarterly*, *68*(2), 221–243.
Wennberg, J. E. (1984). Dealing with medical practice variations: A proposal for action. *Health Affairs (Millwood)*, *3*(2), 6–32.
Wholey, D., Feldman, R., & Christianson, J. B. (1995). The effect of market structure on HMO premiums. *Journal of Health Economics*, *14*, 81–105.
Wrightson, D. W. (1990). *HMO rate setting & financial strategies*. Ann Harbor, MI: Health Administration Press.

APPENDIX A. PRACTICE QUALITY ASSESSMENT SCORE (PQAS)

Over the years, Keystone Health Plan East (KHPE) has developed a comprehensive Quality Management Program. This effort includes careful screening and credentialling of physicians before they are approved to participate in KHPE. It also includes periodic recredentialling of all participating physicians to assure that they continue to meet all criteria.

Elements of KHPE's Quality Management Program include (a) MMSs, (b) comprehensive and responsive grievance procedures, (c) focused medical audits, (d) detailed office record reviews (ORRs), and (e) hospital record reviews for selected patients.

KHPE uses results from its Quality Management Program in its PCP reimbursement system. Therefore, the level of reimbursement for a PCP office will be significantly affected by how the office ranks regarding various quality criteria collected as a part of the Quality Management Program.

A.1. Practice Quality Assessment Score

The quality ranking of a PCP office will be reflected in a Practice Quality Assessment Score (PQAS). PQAS is based on two parameters – ORRs and

member survey results. The number of these parameters will be expanded and refined in coming years as continuous improvements are made in obtaining statistically consistent data for all aspects of the KHPE Quality Management program. The PQAS is a simple sum of these two parameters, as follows:

$$\text{PQAS} = \text{MMS Score} + \text{ORR Score}$$

Each of the variables is weighted equally, with each variable being scored from 0 (lowest) to 4 (highest). Thus, the highest PQAS possible is 8.0.

A.2. Office Record Review

In its ORR process, KHPE uses dedicated and highly trained nurse quality coordinators to perform on-site reviews of a random sample of records. The review encompasses tests for the presence or absence of significant elements, which are appropriate to the delivery of ambulatory care and the development of a health maintenance program for KHPE members.

The test elements include: medical history, biographical information, immunization, smoking, and substance abuse history, problem list, and follow-up, among others.

A.3. Practice Quality Assessment Score

The test elements have been provided to participating physicians and are accepted by professional societies and organizations as standards for ambulatory medical records. Following the review by the nurse quality coordinator, an ORR score is developed using elements found in the chart review. This becomes one of the components of the PQAS.

A.4. Member Satisfaction Survey

The other component used in PQAS is the MSS score. This score is obtained from a survey of subscribers within a primary care practice. KHPE believes that its members are important monitors of the quality of care and the quality of service accessible through the primary care practice. To obtain information from its members about their perceptions of access, quality of service, and quality of care, KHPE uses a mailed survey seeking responses

about waiting times, after hours access, communication with the physician, and services from the office staff. In addition, the member is asked to recommend or not recommend the physician to other KHPE members.

A.5. Future Development

The PQAS will expand in future years to include other components where objective, measurable criteria are available. The focus will be toward the PCP, the office setting, and the member, since all are significant to the growth and development of KHPE. As the system evolves, it will include additional factors that demonstrate a commitment by the practice to improve the quality of care and the quality of services accessible to KHPE members. Such factors will include a focus on preventive health services and early diagnosis and treatment of problems, essential characteristics of a successful health maintenance program.

A.6. Membership Criteria

The PQAS is calculated only for offices with 75 members or more. Below that level, there is an insufficient number of members to calculate a statistically valid PQAS for use in QIPS. However, member surveys, ORRs, medical audits, and the other elements of KHPE's Quality Management Program continue to be conduced for all participating PCP offices.

A.7. Example of Quality Incentive Calculation

This example uses the following assumptions to illustrate how one HMO plan adds quality-based incentives to a basic capitation rate (a fixed payment per member per month) to yield a total payment per member per month.

Assumptions:

- Family practice with an average age group
- 1,000 members
- Accepts existing patients only as new KHPE members
- PQAS score of 7.5 (in top 40% of offices)
- Meets Full Service Office Criteria
- Specialist utilization = 0.2 S.D. worse than KHPE average
- Facility utilization = 0.7 S.D. better than KHPE average

Capitation (paid monthly):

- \$10.23 (floor capitation for an average age group)
- \$0.40 (Full Service Office Premium)
- \$0.80 (High Quality Office Premium) \$11.43 PMPM × 1,000 members = \$11,430

Bimonthly Incentive (paid every other month):

- \$0.50 (based on specialist utilization)
- \$1.00 (based on facility utilization) \$1.50 PMPM × 1.1 Quality Adjustment Factor × 1,000 members × 2 months = \$3,300

Total Payments over a Year (assuming constant performance):

- (\$11,430 × 12) + (\$3,300 × 6) = \$156,960 = \$13.08 PMPM

APPENDIX B

B.1. Area Resource File (ARF) Description

The ARF is a database compiled by the Office of Research and Planning at the Bureau of Health Professions. The data represent a consolidation of sources, including census data and surveys by the American Hospital Association and InterStudy. The time periods covered in this database range from 3 to 15 years depending on the data source. All data are at the county level. For example, the number of doctors per capita is computed for each county in the U.S.

B.2. Calculation of Market Variables

Market variables (DOCCAP, MKTPWR, and BEDS) were constructed from the ARF, which provides an extensive range of demographic information at the county level. Rather than estimate county enrollment for each HMO to construct market area variables, previous research on contract choice simply defined the HMO market area as the county in which the HMO was headquartered. For example, income per capita for an HMO headquartered in Cook County would be obtained by taking the average per capita income for Cook County from the ARF. However, Wholey et al. (1995) argue for including all counties served by the HMO because most

HMOs serve counties other than those in which they are headquartered. Using an approach similar to that of Wholey et al. (1995), we first obtain HMO enrollment data by Metropolitan Statistical Area (MSA) from InterStudy, and then follow the procedure outlined below.[19]

Enrollment for each HMO is first separated into MSA and non-MSA enrollment. MSA enrollment is the number of an HMO's enrollees that reside within the MSAs the HMO is serving. If, for example, an HMO's service area includes two MSAs, total MSA enrollment for the HMO consists of the total of the enrollment within each MSA. Non-MSA enrollment is the number of enrollees that reside outside of an MSA. A weight is then assigned to each MSA as a proportion of the total HMO population. For example, if one quarter of an HMO's enrollment is in the Chicago MSA, a weight of 0.25 is assigned to the Chicago MSA. The counties within an MSA are weighted based on the population of that county relative to the population of the other counties in the MSA. To continue the example, if 80% of the population within the Chicago MSA is in Cook County, the final weighting for Cook County for the HMO will be $0.8 \times 0.25 = 0.2$. The enrollment in Non-MSA counties is weighted based on the population of each county relative to the total Non-MSA population for the HMO. These weightings reflect an estimate of the proportion of enrollees in each county served by the HMO. The calculation of the DOCCAP, MKTPWR, and BEDS variables is then based on a weighted average of the county values, where the weights are determined as described above.

This page intentionally left blank

REVENUE DRIVERS: REVIEWING AND EXTENDING THE ACCOUNTING LITERATURE

Jeffrey F. Shields and Michael D. Shields

ABSTRACT

While management-accounting research continues to focus on cost drivers, research has recently begun to examine revenue drivers. We review the research on revenue drivers with reference to five revenue-driver models in the accounting literature. The revenue drivers identified by quantitative empirical research are located in a revenue-driver model based on their levels of analysis (customer, product, organization, industry) and other characteristics of a revenue driver–revenue relation. Implications of this model for research are discussed.

1. INTRODUCTION

During the last 20 years, several profit- (or value-) driver models have been developed (Porter, 1985; Koller, 1994; Foster, Gupta, & Sjoblom, 1996; Kaplan & Norton, 1996, 2004; Epstein, Kumar, & Westbrook, 2000; Ittner & Larcker, 2001). While these models are intended to focus on profit, most models emphasize cost over revenue. For example, in the early 1990s, two

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 33–60

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14002-7

cost-driver models were developed: a hierarchy of activity-based cost (ABC) drivers (Cooper, 1990) and a set of strategic structural and executional cost drivers (Shank & Govindarajan, 1992). While these cost-driver models have provided a lot of insight into explaining and predicting costs, they provide an incomplete understanding of profit drivers, including revenue drivers.

The purpose of this paper is to increase understanding of revenue drivers through several steps in the ensuing sections of this paper. Section 2 defines a revenue driver and identifies characteristics of a revenue driver–revenue relation. Section 3 reviews and compares existing revenue-driver models in the accounting literature and identifies revenue drivers supported by quantitative evidence that are related to these revenue-driver models. In Section 4, we locate these revenue drivers in a model based on their levels of analysis (customer, product, organization, and industry) and other characteristics of revenue driver–revenue relations. Section 5 discusses implications of this model for research.

2. REVENUE DRIVER–REVENUE RELATIONS

A revenue driver is defined as a variable that influences revenue (Horngren, Datar, & Foster, 2006). The relation between a revenue driver and revenue can be described by several characteristics of the relation. A relation's causal-model form specifies how an independent variable (revenue driver) causally influences a dependent variable (revenue), including additivity (additive or interactive), linearity (linear or curvilinear), directness (direct or indirect), and direction (unidirectional or bidirectional causality and if bidirectional nonrecursive reciprocal or reciprocal nonrecursive) (Luft & Shields, 2003). Other characteristics include level of analysis (customer, product, organization, industry), sign (positive or negative), timing (contemporaneous or lead), and duration (temporal length of effect) (Kennedy, 2003; Luft & Shields, 2003; Hanssens, Parsons, & Schultz, 2001). The remainder of this section describes each of these characteristics of revenue driver–revenue relations that are or should be addressed by accounting research. It also provides a guide to interpreting the analysis, evidence, and models presented later.

2.1. Level of Analysis

Level of analysis of a variable is the level at which its variation of interest occurs (Luft & Shields, 2003). For research on revenue drivers, pertinent

levels of analysis include customers, products, organizations, and industries. A single-level study at, for example, the customer level assumes that variation in revenue and its driver(s) occurs between customers; thus attributes of customers are revenue drivers (e.g., education, preferences, wealth) and revenue is for each customer. In contrast, a study at the organization level assumes variation in the variables that occurs between organizations (e.g., organization strategy). A study should have alignment among its level of the theory, level of variable measurement, and level of data analysis. In the revenue-driver studies to be reviewed, their theory, variable measurement, and data analysis most frequently are at the organizational level.

2.2. Sign

The relation between a revenue driver and revenue can be positive or negative. For example, web usage can have a positive effect on revenue (Trueman, Wong, & Zhang, 2001). In contrast, low-quality products, defined as defect rates, can have a negative effect on revenue (Nagar & Rajan, 2001).

2.3. Linearity

A revenue driver can have linear effects on revenue, which means that its effect on revenues is constant over all of its values (e.g., a unit increase in a revenue driver always increases revenue by \$10). In contrast, as suggested by economic theory, the effects of a revenue driver can be curvilinear; for example, a relation can capture diminishing or increasing returns such that the effect of a revenue driver on revenue decreases or increases as its value increases. A linear relation can also be a step-function. For example, the relation between customer satisfaction and revenue can vary within a range of customer satisfaction: for low to medium levels of customer satisfaction, revenue might increase \$15 with each unit of increase in customer satisfaction, and for medium to high levels of customer satisfaction, revenue might increase by \$10 per unit increase in customer satisfaction. Ittner and Larcker (1998b) provide evidence that customer satisfaction can have a linear, a step-function linear, and/or a curvilinear relation with revenue.

2.4. Additivity

A revenue driver can have an additive or interactive effect on revenue. An additive effect means that the effect of a revenue driver on revenue is

constant across levels of other variables. For example, customer satisfaction has additive effects on revenue (Ittner & Larcker, 1998b), meaning that its effect on revenue is not influenced by the value of other variables.

In contrast, the effects of some revenue drivers on revenue are interactive; that is, they are conditional on the value of another variable. For example, in banks, the impact of volume (e.g., number of accounts, number of new accounts) on revenue depends on how they use information technology. Banks that use information technology to increase market share find that volume has a larger effect on their revenue than banks that do not use technology to increase market share (Mistry & Johnston, 2004). For ordinal interactions such as this one, the sign of revenue driver–revenue relation is constant but its magnitude depends on the value of another variable. For a disordinal interaction, the sign (and usually the magnitude) of a revenue–driver relation depends on the value of another variable. For example, evidence from online retailing indicates that website stickiness (e.g., time spent on a website) has positive effects on the percent change in next quarter's return on revenue when there is a high level of website satisfaction; but when there is a low level of website satisfaction, stickiness has negative effects on the change in next quarter's return on revenues (Dikolli & Sedatole, 2003).

2.5. Directness

Some revenue drivers have direct effects on revenue while others affect revenue indirectly by affecting another revenue driver that does directly affect revenue. Some examples of direct and indirect revenue drivers are given below. First, on-time delivery and defect rates can have direct effects on revenue (Nagar & Rajan, 2001). Second, an example of an indirect effect is Sears' causal business model that predicts that a five-unit increase in employee attitude will cause a 1.3 unit increase in customer satisfaction, which in turn will increase revenue by 0.5% (Rucci, Kirn, & Quinn, 1998). Third, post-sales service quality can have indirect effects on revenue through its effects on customer loyalty (Smith & Wright, 2004). Smith and Wright (2004) also provide evidence that customer loyalty can have indirect effects on revenue through its effect on price. Last, incentive plans have been shown to affect customer satisfaction, which then affects revenue (Banker, Potter, & Srinivasan, 2000b).

2.6. Directionality

When a revenue driver influences revenue and revenue does not influence the revenue driver, the relation is unidirectional causality (Luft & Shields, 2003).

A relation can also have bidirectional causality; for example, a revenue driver such as sale-force design can influence revenue, and revenue can influence sales-force design. In bidirectional relations, a cyclical recursive relation has an identifiable time interval between the change in one variable and the corresponding change in the other variable and vice versa. In contrast, in a reciprocal nonrecursive relation, the two variables are determined simultaneously or at intervals too short for the causal influences in different directions to be distinguished empirically. All of the accounting studies on revenue drivers assume unidirectional causation, therefore, this characteristic of revenue driver–revenue relations is not incorporated into the review of the extant accounting literature.

2.7. Timing

The effects of some revenue drivers on revenue are contemporaneous (or in accounting terms, within the same measurement and reporting period). For example, customer satisfaction can have contemporaneous effects on revenue in the airline industry (Behn & Riley, 1999). In contrast, the effects of some revenue drivers on revenue are leading (or in accounting terms, in future periods). For example, product quality, defined as internal and external failure costs, defect rates, and on-time delivery, can affect sales with one-quarter, two-quarter, three-quarter, and four quarter leads (Nagar & Rajan, 2001).

2.8. Duration

Duration is the length of time a revenue driver has an effect on revenue. For example, marketing research indicates that the duration of the effect of advertising on revenue is 6–9 months (Assmus, Farley, & Lehmann, 1984; Leone, 1995). The quantitative empirical accounting research on revenue drivers does not address duration, so this characteristic is not incorporated into the review of the extant accounting literature.

2.9. Guide to Review of Accounting Studies

In reviewing the accounting literature, we present studies that provide analysis, qualitative evidence (e.g., case study), and/or quantitative evidence

(e.g., econometric analysis of archival data) on revenue drivers.[1] The quantitative empirical research for each study is summarized in Table 1. The information in Table 1 is based on the reported results of studies for their non-control variables. We limit the quantitative evidence reviewed to studies that have as their dependent variable sales revenue in a regression analysis or a structural equation model. These revenue measures vary in whether they are scaled; for example, some studies use sales divided by assets or square foot. For expositional convenience, the text will refer to revenue regardless as to how it is measured; the tables report the actual revenue measure used. Furthermore, we assume that revenue driver–revenue relations in quantitative empirical studies are positive linear additive direct contemporaneous, because it is by far the most frequently identified relation. Therefore, for parsimony for the quantitative empirical studies, only those characteristics of a revenue driver–revenue relation that deviate from these assumed characteristics are included in the text. The revenue driver–revenue relations described in studies that report qualitative evidence and/or analysis are presented in the text.

3. REVENUE-DRIVER MODELS IN THE ACCOUNTING LITERATURE

The following review includes five revenue-driver models in the accounting literature. These models vary in their components (performance measurement, causal business model, and strategic management system) and in how they assume revenue is influenced. The models are ordered approximately by their entry into the management accounting literature and differ in their emphasis on performance measurement, causal business models, and strategic management systems.

3.1. Activity-Based Costing

This model focuses on measuring the supply and consumption of overhead resource costs by cost objects such as products and customers by the drivers of activity costs (Cooper, 1990; Kaplan & Cooper, 1998). ABC assumes that several types of activity cost drivers are structured as a non-causal hierarchy of drivers, including unit, batch, product, facility, and customer. Past quantitative research has focused on identifying cost drivers and their relations with costs; while more recent research has begun to examine the effects of

Table 1. Revenue-Driver Models in the Accounting Literature: Quantitative Evidence.

Study Number and Study	Dependent Variable[a]	Independent Variable[a]	Sign	Linearity	Additivity	Directness[b]	Timing
Panel A: Activity-Based Costing							
1. Ittner, Larcker, and Randall (1997)	Revenue (P)	Product-sustaining activities (P)	Positive	Linear	Additive	Direct	Contemporaneous
1. Ittner et al. (1997)	Revenue (P)	Unit production volume (P)	Positive	Linear	Additive	Direct	Contemporaneous
2. Mistry and Johnston (2004)	Revenue (O)	Volume (O) × strategic information technology (O)	Positive	Linear	Ordinal interaction	Direct	Contemporaneous
Panel B: Strategic Cost Analysis							
3. Banker, Chang, and Cunningham (2003)	Revenue (O)	Economies of scale (O)	Positive	Curvilinear	Additive	Direct	Contemporaneous
4. Petersen (2003)	Revenue (O)	Competition (O)	Negative	Linear	Additive	Direct	Contemporaneous
4. Petersen (2003)	Revenue (O)	Operating scope (O)	Negative	Linear	Additive	Direct	Contemporaneous
4. Peterson (2003)	Revenue (O)	Physical distance between locations of initial and subsequent customer contact (O)	Negative	Linear	Additive	Direct	Contemporaneous
Panel C: Balanced Scorecard							
5. Banker, Konstans, and Mashruwala (2000a)	Revenue (O)	Customer satisfaction (O)	Positive	Linear	Additive	Direct	Contemporaneous
5. Banker et al. (2000a)	Revenue (O)	Employee satisfaction (O)	Positive	Linear	Additive	Direct	Contemporaneous
6. Bryant, Jones, and Widener (2004)	Revenue (O)	Market share (O)	Positive	Linear	Additive	Direct	Contemporaneous

Table 1. (*Continued*)

Study Number and Study	Dependent Variable[a]	Independent Variable[a]	Sign	Linearity	Additivity	Directness[b]	Timing
6. Bryant et al. (2004)	Revenue (O)	Customer satisfaction (O)	Positive	Linear	Additive	Indirect: market share	Contemporaneous
6. Bryant et al. (2004)	Revenue (O)	Employee skill (O)	Positive	Linear	Additive	Indirect: customer satisfaction, market share	Contemporaneous
6. Bryant et al. (2004)	Revenue (O)	New product and service introduction (O)	Positive	Linear	Additive	Indirect: customer satisfaction, market share	Contemporaneous
Panel D: Nonfinancial Performance Measures							
7. Banker and Mashruwala (2001)	Sales per square foot (O)	Customer satisfaction (O)	Positive	Linear	Additive	Direct	Contemporaneous
8. Banker et al. (2000b)	Revenue per available room (O)	Customer satisfaction (O)	Positive	Linear	Additive	Direct	Mean lead of 3 months
8. Banker et al. (2000b)	Revenue per available room (O)	Employee incentives (O)	Positive	Linear	Additive	Indirect: customer satisfaction	Contemporaneous
9. Behn and Riley (1999)	Operating revenue (O)	Capacity (O)	Positive	Linear	Additive	Direct	Contemporaneous
9. Behn and Riley (1999)	Operating revenue (O)	Capacity utilization (O)	Positive	Linear	Additive	Direct	Contemporaneous
9. Behn and Riley (1999)	Operating revenue (O)	Customer satisfaction (O)	Positive	Linear	Additive	Direct	Contemporaneous

9. Behn and Riley (1999)	Operating revenue (O)	Market share (O)	Positive	Linear	Additive	Direct	Contemporaneous
10. Dikolli and Sedatole (2003)	Return on revenue (O)	Website stickiness (O) × website satisfaction (O)	Return on revenue – .website stickiness relation is positive (negative) when website satisfaction is high (low)	Linear	Disordinal interaction	Direct	1-quarter lead
11. Dikolli, Kinney, and Sedatole (2004)	Revenue (O)	Customer satisfaction (O) × switching Costs (O) × pure internet firm (O)	Positive	Linear	Ordinal interaction	Direct	1-quarter lead
12. Foster and Gupta (1999)	Change in sales volume (C)	Change in customer satisfaction (C)	Positive	Linear	Additive	Direct	Contemporaneous
13. Ittner and Larcker (1998b)	Revenue (C)	Customer satisfaction (C)	Positive	Linear	Additive	Direct	1-year lead
13. Ittner and Larcker (1998b)	Change in revenue (C)	Customer satisfaction (C)	Positive	Curvilinear	Additive	Direct	1-year lead
13. Ittner and Larcker (1998b)	Revenue (O)	Customer satisfaction (O)	Positive	Linear	Additive	Direct	2-quarter lead
13. Ittner and Larcker (1998b)	Revenue (O)	Customer satisfaction (O)	Positive	Curvilinear	Additive	Direct	1-year lead
14. Ittner, Larcker, and Randall (2003)	3-year sales growth (O)	Nonfinancial performance measures for the value drivers (actual − expected use) (O)	Negative	Linear	Additive	Direct	3-year lead

Table 1. (*Continued*)

Study Number and Study	Dependent Variable[a]	Independent Variable[a]	Sign	Linearity	Additivity	Directness[b]	Timing
14. Ittner et al. (2003)	3-year sales growth (O)	Performance measures for the competitive strategy (actual − expected use) (O)	Negative	Linear	Additive	Direct	3-year lead
15. Nagar and Rajan (2001)	Percent change in sales (O)	Defect rates (O)	Negative	Linear	Additive	Direct	Leads of 1 and 4 quarters
15. Nagar and Rajan (2001)	Percent change in sales (O)	On-time delivery (O)	Positive	Linear	Additive	Direct	1-quarter lead
16. Smith and Wright (2004)	Sales growth (O)	Customer loyalty (O)	Positive	Linear	Additive	Direct	Contemporaneous
16. Smith and Wright (2004)	Sales growth (O)	Firm viability (O)	Positive	Linear	Additive	Direct	Contemporaneous
16. Smith and Wright (2004)	Sales growth (O)	Price (O)	Positive	Linear	Additive	Direct	Contemporaneous
16. Smith and Wright (2004)	Sales growth (O)	Customer loyalty (O)	Positive	Linear	Additive	Indirect: product price	Contemporaneous
16. Smith and Wright (2004)	Sales growth (O)	Brand (O)	Positive	Linear	Additive	Indirect: customer loyalty	Contemporaneous
16. Smith and Wright (2004)	Sales growth (O)	Firm viability (O)	Positive	Linear	Additive	Indirect: customer loyalty	Contemporaneous
16. Smith and Wright (2004)	Sales growth (O)	Post-sale service quality (O)	Positive	Linear	Additive	Indirect: customer loyalty	Contemporaneous

16. Smith and Wright (2004)	Sales growth (O)	Product quality (O)	Negative	Linear	Additive	Indirect: customer loyalty	Contemporaneous
17. Trueman et al. (2001)	Revenue growth (O)	Web traffic growth (unique visitors, page views) (O)	Positive	Linear	Additive	Direct	Contemporaneous

[a]Level of analysis: (C) = Customer; (P) = Product; (O) = Organization.
[b]Directness with respect to revenue. Mediating variables are listed after the indirect notation.

cost drivers on revenues. The basis for using the ABC model to explain revenues is that some activities have been found to influence revenues (Ittner et al., 1997). Unit activities (such as production volume) increase revenue. Besides unit drivers, theoretical and empirical studies in marketing indicate that product-sustaining activities that affect the breath of product lines can increase or decrease revenues, depending on whether the change in product variety increases or decreases the product brand image or fragments-marketing efforts (see Ittner et al., 1997).

Quantitative evidence shows that some activities are revenue drivers (Table 1, Panel A). This evidence is at the product and organization levels of analysis. Ittner et al. (1997) provide evidence that unit- and product-related activity cost drivers affect revenue. Mistry and Johnston (2004) found evidence in banking that a unit (volume) cost driver has an ordinal interaction effect with strategic information technology spending (e.g., ATMs, home banking) on revenue. Qualitative evidence indicates that new-customer-screening and cross-selling activities also have positive effects on revenue (Kaplan & Narayanan, 2001).

3.2. Strategic Cost Analysis

This model is intended to provide performance measurement of how an organization's competitive advantage is influenced by its strategic structural and executional cost drivers (Shank & Govindarajan, 1989, 1992). Shank and Govindarajan (1989) provide qualitative analysis of the 12 structural and executional variables that Porter (1985) assumes drive costs, based on his qualitative analysis of the industrial-organization economics' literature (Scherer & Ross, 1990). These 12 cost drivers are discretionary policies, economies of scale, interrelationships among business units, institutional factors, learning, level of vertical integration, links with channels, links with suppliers, links within the value chain, location, pattern of capacity utilization, and timing. Within this economics literature, much research has been based on the structure-conduct-performance model, which predicts that the structure of an industry determines the opportunities and constraints for organizations in that industry (Bain, 1968; Scherer & Ross, 1990). These opportunities and constraints determine organizations' conduct (e.g., capital investments, pricing, strategy) and hence their performance (e.g., efficiency, profit) (Barney, 1997). The evidence that addresses the structure-conduct-performance model uses measures of profit or profitability as the dependent variable.

Porter (1980, 1985) and Shank and Govindarajan (1989, 1992) assume that these drivers of profit also are drivers of cost. Some accounting studies based on the strategic cost analysis model, likewise, assume that these cost drivers are drivers of revenue, excluding the pattern of capacity utilization, because as drivers of organizational differentiation (Porter, 1985) they create customer value and thus drive revenue.

Evidence based on this model is at the organization level of analysis. Banker et al. (2003) provide quantitative evidence that the largest 100 public accounting firms' production function (skill level of employees engaged in various value chain activities) has a curvilinear effect on revenue with increasing returns (i.e., economies of scale) (Table 1, Panel B). The production function of a company is a discretionary policy, which Porter (1985, p. 124) claims determine "...what activities to perform and how to perform them." Petersen (2003) reports quantitative evidence on downstream revenue drivers, which are defined as revenue that is not earned immediately at the initial customer contact but which arises from subsequent contact with the customer (e.g., post-purchase services). He finds that three structural revenue drivers have negative effects on downstream revenue: competition, operating scope, and physical distance between the initial and later customer contact.

3.3. Balanced Scorecard

This model consists of three parts: performance measurement, strategy map, and strategic management process. Performance measurement is an integrated set of performance measures that are structured to form a causal hierarchy with four levels called perspectives – learning and growth, internal business processes, customer, and financial – that are related to an organization's competitive strategy (Kaplan & Norton, 1996). The strategy map is a causal business model that specifies how the learning and growth, internal business processes, and customer perspectives are causally linked in the visual form of a map, which indicates how the various performance measures are causally linked in paths that typically end at financial performance – including revenues – as the terminal dependent variable (Kaplan & Norton, 2004). The strategic management system is comprised of five principles, – translating strategy into operational terms, align the organization to the strategy, make strategy everyone's everyday job, make strategy a continual process, and mobilize change through executive leadership – which are intended to make organization's strategy focused (Kaplan & Norton, 2001).

Kaplan and Norton (1996, 2001, 2004) present qualitative analysis and evidence to support their assumption that the four perspectives form a causal hierarchy in which the outcomes at one level of the hierarchy are direct drivers of the next higher level in the hierarchy and indirect drivers of the other higher levels in the hierarchy. For example, employee training increases service quality which then increases customer satisfaction, which in turn increases customer loyalty which then finally increases revenues. Kaplan and Norton (2004) assume that customer and internal processes create value for customers by differentiating products, which then drives revenue.

Several studies use the balanced scorecard to structure their evidence on revenue drivers. Kaplan and Norton (2001) present qualitative evidence that the following are revenue drivers: product price, product mix, product variety, new customers, new products/services, the shopping experience, and integrated services across value-chain activities, but they do not specify the exact characteristics of the relation (e.g., sign, additivity, duration).

Some quantitative empirical studies use the balanced scorecard as the basis for organizing their evidence (Table 1, Panel C). These studies present evidence at the organization level of analysis. Banker et al. (2000a) find that customer satisfaction and employee satisfaction affect revenue. Bryant et al. (2004) report that market share directly affects revenue. They show that customer satisfaction has indirect effects on revenue meditated by market share. Lastly, Bryant et al. (2004) provide evidence that employee skill and new product/service introduction have indirect effects on revenue mediated by customer satisfaction and market share.

3.4. Nonfinancial Performance Measures

This measurement model identifies nonfinancial performance measures that are drivers of organizations' economic value (e.g., revenue) (Ittner & Larcker, 1998b). These measures are assumed to be more informative, timely, or actionable than nonfinancial performance measures (Ittner & Larcker, 1998a). They are also assumed to have earlier and higher ability to predict future financial performance, including revenues, than can financial measures. The assumption is that some nonfinancial performance measures' referent (the object of the measurement such as customer satisfaction or product quality) influence revenue, particularly revenue in later periods. The quantitative evidence based on this model is predominately at the organization level, although it does provide the only evidence from all of the revenue-driver models that is at the customer level of analysis.

Several studies provide quantitative evidence that customer satisfaction is a revenue driver but that the relation can have various characteristics (Table 1, Panel D). For example, some evidence indicates that customer satisfaction measured by customers' responses to questions can have linear and/or curvilinear effects on revenue (Ittner & Larcker, 1998b; Foster & Gupta, 1999). This relation can have various timing effects with time leads of one year (Ittner & Larcker, 1998b), two-quarters (Ittner & Larcker, 1998b), one-quarter (Banker et al., 2000a), and contemporaneous (Banker & Mashruwala, 2001; Behn & Riley, 1999; Foster & Gupta, 1999). This relation can also be interactive: for example, website stickiness (e.g., time spent on a website) and website satisfaction have a disordinal interaction effect on revenues (Dikolli & Sedatole, 2003).

Besides customer satisfaction, other nonfinancial performance measures can influence revenue. Nagar and Rajan (2001) show that defect rates and on-time performance are revenue drivers. In particular, they provide evidence that the effect of defect rates on revenue is negative with one- and four-quarter leads and that on-time delivery has a one-quarter lead effect on revenue. Smith and Wright (2004) provide evidence that customer loyalty, firm viability (i.e., going concern), and product price have effects on revenue. They also show that customer loyalty has an indirect effect on revenue mediated by product price. Lastly, brand, firm viability, post-sales service quality, and product quality have indirect effects on revenue meditated by customer loyalty with product quality having a negative effect on revenue. Trueman et al. (2001) report that web traffic growth (unique visitors, page views, and minutes) affects revenue. Finally, Ittner et al. (2003) find that organizations' performance measurement system can influence revenue. They report evidence of negative relations with 3-year leads between revenue and the extent to which organizations use performance measures more than expected given their competitive strategy or value drivers.

3.5. Action-Profit-Linkage

This causal-business model has a long, complex causal chain to explain revenues, beginning with organization actions (e.g., marketing providing customer service, human resources providing employee training), which influence the delivered product, which in turn affects customer actions, finally influencing revenues (Epstein et al., 2000; Epstein & Westbrook, 2001). It assumes that an organization's actions, delivered product, and customer actions are revenue drivers. The key part of this model is customer actions,

which is comprised of three variables each of which have subcategory variables: the delivered product is assumed to affect (1) customer perceptions of the product which in turn affects (2) customers' attitudes (e.g., customer satisfaction and loyalty) and then (3) customer behavior (e.g., initial purchase, repeat purchase, related purchases, new customer referrals, and price acceptance). At present, in the accounting literature, there is no quantitative evidence on the validity of this causal business model.

3.6. Comparison of Models

The five models reviewed can be compared with respect to their main components – performance measurement, causal business models, and strategic management system – and their link to revenue. Starting with the balanced-scorecard model because it is the most complete or complex model, it has three components: financial and nonfinancial performance measures, causal business model (strategy map), and a strategic management system. The action-profit-linkage model is a causal business model. The ABC, strategic cost analysis, and nonfinancial performance measures models focus on performance measurement, measurement of the cost of activities, measurement of the cost of structural and executional cost drivers, and nonfinancial performance measurement, respectively.

These five models also vary in their assumptions about revenue driver–revenue relations. The ABC model assumes that cost drivers influence revenue via unit and product-sustaining activities, with the latter affecting revenue by the breath of the product line or product variety. The strategic cost analysis model assumes that structural and executional cost drivers influence revenue because as drivers of organizational differentiation they create customer value and thus drive revenue. The balanced scorecard model assumes that revenues are influenced by customer and internal processes. The nonfinancial performance measures model assumes that some nonfinancial performance measures' referent (e.g., customer satisfaction, product quality) influence future revenue. Finally, the action-profit-linkage model assumes that an organization's actions, delivered product, and customer actions influence revenue.

4. A LEVEL-OF-ANALYSIS MODEL

The five proposed revenue-driver models and their revenue drivers in Table 1 provide a lot of analysis and/or evidence on revenue driver–revenue

relations. This information, however, is scattered because the various models have different foci (e.g., their relative or direct emphasis on revenues, costs, and profits), levels of analysis, and diversity of variables and data in the empirical studies, which limits our overall understanding of revenue drivers from these studies.

To provide a more complete and systematic understanding of revenue drivers, we organize the revenue drivers in Table 1 based on the levels of analysis of their independent variables. We use levels of analysis to locate the revenue drivers in the model because a variable's level of analysis is expected to be based on theory, variable measurement, and data analysis, all at the same level (Luft & Shields, 2003). Every revenue driver–revenue relation (row) in Table 1 is located in Table 2 by reference to the level of analysis of its dependent and independent variables because all variables in these relations are at the same level of analysis.

The model has four levels of analysis. Customer level revenue drivers are attributes of customers that differ between customers and affect revenue through characteristics of customers (e.g., customer satisfaction). Product revenue drivers are attributes of products that vary between products and these attributes influence revenue (e.g., price, quality). Organizational revenue drivers cause changes in revenues from the use of organizations' resources and/or the actions of employees that affect revenues (e.g., competitive strategy, product development). Industry revenue drivers are characteristics of an industry that affect competitive forces and thus the price and volume of products sold in an industry (e.g., competition, economies of scale, regulations).

At the customer level (Table 2, Panel A), the only revenue-driver research in the accounting literature is customer satisfaction. While this evidence has similarities across the three customer-level bivariate relations (same independent and dependent variables, positive, additive direct), there are differences (linearity, timing), and no evidence on duration or bidirectionality.

At the product level (Table 2, Panel B), the revenue drivers are a set of related product-sustaining activities (e.g., purchase orders, orders shipped) and product volume. Both product-level drivers have the same relation with revenue – positive, linear, additive, direct, contemporaneous – but provide no evidence on duration or bidirectionality.

At the organization level (Table 2, Panel C), 25 revenue drivers have been identified. These drivers have a variety of relations with revenue, but the most common relation is positive, linear, additive, direct, and contemporaneous, but with no evidence on duration or bidirectionality.

Table 2. Revenue-Driver Model Based on Levels of Analysis.

Revenue Driver	Study Number[a]	Sign	Linearity	Additivity	Directness[b]	Timing
Panel A: Customer-Level Revenue Drivers						
Customer satisfaction	12	Positive	Linear	Additive	Direct	Contemporaneous
Customer satisfaction	13	Positive	Linear	Additive	Direct	1-year lead
Customer satisfaction	13	Positive	Curvilinear	Additive	Direct	1-year lead
Panel B: Product-Level Revenue Drivers						
Product-sustaining activities	1	Positive	Linear	Additive	Direct	Contemporaneous
Unit production volume	1	Positive	Linear	Additive	Direct	Contemporaneous
Panel C: Organization-Level Revenue Drivers						
Brand	16	Positive	Linear	Additive	Indirect: customer loyalty	Contemporaneous
Capacity	9	Positive	Linear	Additive	Direct	Contemporaneous
Capacity utilization	9	Positive	Linear	Additive	Direct	Contemporaneous
Competition	4	Negative	Linear	Additive	Direct	Contemporaneous
Customer loyalty	16	Positive	Linear	Additive	Direct	Contemporaneous
Customer loyalty	16	Positive	Linear	Additive	Indirect: product price	Contemporaneous
Customer satisfaction	7	Positive	Linear	Additive	Direct	Contemporaneous
Customer satisfaction	8	Positive	Linear	Additive	Direct	Mean lead of 3 months
Customer satisfaction	9	Positive	Linear	Additive	Direct	Contemporaneous
Customer satisfaction × switching costs × pure internet firm	11	Positive	Linear	Ordinal interaction	Direct	1-quarter lead
Customer satisfaction	13	Positive	Linear	Additive	Direct	2-quarter lead
Customer satisfaction	13	Positive	Curvilinear	Additive	Direct	1-year lead
Customer satisfaction	5	Positive	Linear	Additive	Direct	Contemporaneous
Customer satisfaction	6	Positive	Linear	Additive	Indirect: market share	Contemporaneous
Defect rates	15	Negative	Linear	Additive	Direct	Leads of 1 and 4 quarters

Economies of scale	3	Positive	Curvilinear	Additive	Direct	Contemporaneous
Employee incentives	8	Positive	Linear	Additive	Indirect: customer satisfaction	Contemporaneous
Employee satisfaction	5	Positive	Linear	Additive	Direct	Contemporaneous
Employee skill	6	Positive	Linear	Additive	Indirect: customer satisfaction, market share	Contemporaneous
Firm viability	16	Positive	Linear	Additive	Direct	Contemporaneous
Firm viability	16	Positive	Linear	Additive	Indirect: customer loyalty	Contemporaneous
Market share	9	Positive	Linear	Additive	Direct	Contemporaneous
Market share	6	Positive	Linear	Additive	Direct	Contemporaneous
New product and service introduction	6	Positive	Linear	Additive	Indirect: customer satisfaction, market share	Contemporaneous
Nonfinancial performance measures for the value drivers (actual − expected use)	14	Negative	Linear	Additive	Direct	3-year lead
On-time delivery	15	Positive	Linear	Additive	Direct	1-quarter lead
Operating scope	4	Negative	Linear	Additive	Direct	Contemporaneous
Performance measures for the competitive strategy (actual − expected use)	14	Negative	Linear	Additive	Direct	3-year lead
Physical distance between locations of initial and subsequent customer contact	4	Negative	Linear	Additive	Direct	Contemporaneous
Post-sale service quality	16	Positive	Linear	Additive	Indirect: customer loyalty	Contemporaneous
Product price	16	Positive	Linear	Additive	Direct	Contemporaneous
Product quality	16	Negative	Linear	Additive	Indirect: customer loyalty	Contemporaneous

Table 2. (*Continued*)

Revenue Driver	Study Number[a]	Sign	Linearity	Additivity	Directness[b]	Timing
Volume × strategic information technology	2	Positive	Linear	Ordinal Interaction	Direct	Contemporaneous
Web traffic growth (unique visitors, page views)	17	Positive	Linear	Additive	Direct	Contemporaneous
Website stickiness × website satisfaction	10	Return on revenue – website stickiness relation is positive (negative) when website		satisfaction is high (low)	Linear	Disordinal interaction
Direct	1-quarter lead					

[a]See Table 1 for linking of study numbers to studies.
[b]Directness with respect to revenue. Mediating variables are listed after the indirect notation.

The model logically includes an industry level. However, there is no accounting research on revenue drivers at this level.

5. IMPLICATIONS FOR RESEARCH

Our model based on levels of analysis and secondarily on other characteristics of revenue driver–revenue relations (sign, linearity, additivity, directness, timing, direction, and duration) indicates that while a lot of analysis and evidence on revenue drivers exists, much more research remains. The paucity of research is most evident when looking at industry-level drivers in which there is no accounting research. For customer- and product-level drivers there is also a dearth of quantitative evidence with only a couple of drivers identified. Why the predominance of evidence at the organization level compared to the other levels? This could be due to several factors, including: a convenience artifact of accounting data being more available at the organizational level; accounting researchers overly focusing on the organization level because of accounting's focus on organizations; lack of theory at these other levels to guide empirical research; or, most, or the most important, revenue drivers are at the organization level.

5.1. Single-Level Models

Since most accounting research has been at the organization level, future research could put more emphasis on revenue drivers and revenue at the customer, product, and industry levels. A key to such research is to identify and use theory at these various levels in order to guide the selection and measurement of variables and data analysis of revenue driver–revenue relations. The following discussion is illustrative in its identification of variables at different levels of analysis from non-accounting literature that can be used to motivate accounting research.

At the customer level, an extensive search of the marketing literature found no research at the customer level. Discussions with marketing professors indicated that research in marketing is mostly focused on product-level revenue drivers. However, in the economics literature there is some customer-level research based on census or specialized household survey data.

At the product level, the marketing literature provides a lot of quantitative evidence on revenue drivers (Hanssens et al., 2001).[2] Hanssens et al.

(2001) review of this literature identifies five major product-level revenue drivers: advertising, distribution, price, promotion, and sale-force effort. Product advertising is a revenue driver with positive effects on revenue with a duration ranging from 6 to 9 months, but this relation is conditional on brand and market maturity (Sethuraman & Tellis, 1991; Leone, 1995; Shankar, Carpenter, & Krishnamurthi, 1999). In addition, product adverting can have nonlinear effects on revenue (Simon, 1982). Distribution has a bidirectional relation with revenue with a lead of several weeks and this relation is conditional on product life cycle (Bronnenberg, Mahajan, & Vanhonacker, 2000; Reibstein & Ferris, 1995). Meta-analysis of many studies indicates that product price (as well as the price of substitute products) has positive effects on revenue (Tellis, 1988; Sethuraman & Tellis, 1991; Sethuraman, Srinivasan, & Kim, 1999). Product promotions such as coupons, temporary price reductions, and in-store displays influence revenue (Bucklin & Gupta, 1992; Blattberg & Wisniewski, 1989; Kopalle, Mela, & March, 1999). Sales-force effort devoted to a product has a positive nonlinear relation with revenue and can interact with product specialization and sales-force design (Ramaswamy, Desarbo, Reibstein, & Robinson, 1993; Rangaswamy, Sinha, & Zoltners, 1990; Hanssens et al., 2001).

At the organization level, while the preponderance of the accounting research on revenue drivers is at the organization level, other research finds additional organization-level revenue drivers as well some research finding more complex revenue driver–revenue relations. Batt (2002) finds that work design and worker incentives are revenue drivers. Narasimhan and Kim (2002) predict and find that marketing and operations strategy variables (product and market diversification, links to customers, links to suppliers, and internal integration across supply chain) influence revenue with nonlinear interaction effects. As a final example, Babakus, Bienstock, and Van Scottter (2004) predict and find that a company's quality relative to its competitors' quality influences customer satisfaction which in turn influences revenue growth.

At the industry level, the industrial-organization economics literature provides some analysis and evidence on a variant of return on sales, specifically the price-marginal cost margin, which is defined as price minus marginal cost divided by price (Martin, 2002; Scherer & Ross, 1990). In terms of accounting performance, most of this literature is focused on profit and profitability to study the effects of market structure (e.g., monopoly) on economic behavior, but not the components of profit or profitability such as revenue. The focus on the price-marginal cost margin began with the start of industrial-organization economics (Bain, 1968), continuing through the

1970s and 1980s based on the structure–conduct–performance model (Scherer & Ross, 1990), and up to today based on game theory (Martin, 2002). Several studies provide industry-level evidence that the price-marginal cost margin is related to seller concentration, capital-sales ratio, advertising-sales ratio, capacity utilization, industry growth rate, industry elasticity of demand, firm's market share, barriers to entry, and the interaction of firm size and economics of scale (cost advantages) (Martin, 2002; Scherer & Ross, 1990). These drivers of return on sales as revenue drivers at the industry level combined with revenue at the organization level means that most management accounting studies using these industry-level revenue drivers will have cross-level, not single-level, models as discussed in Section 5.2.

5.2. Cross-Level Models

Besides analysis and evidence at each of the four levels (customer, product, organization, and industry), revenue driver–revenue relations can also be cross-level. For example, future research might use a bottom-up model to investigate the effects of product-level variables on organizational-level revenue.

Valid cross-level models require an interaction involving an independent variable(s) (revenue driver(s)) and another independent or a moderator variable(s) at the level of the dependent variable (Luft & Shields, 2003). For example, a top-down model might represent how an industry-level variable such as industry competition (i.e., competition that varies between industries) influences organization-level revenue (revenue that varies between organizations). For industry competition to have different effects on different organizations' revenue requires that organizations differentially respond to industry competition in ways that differentially affect revenue between organizations. For example, firms with different competitive strategies or sale-force designs might be expected to respond to industry-level competition differently, thus differentially influencing their revenues.

Cross-level models can have direct and/or indirect effects of revenue drivers on revenue. For example, consider the following top-down intervening-variable model. At the industry level, at least some of Porter's five-forces (e.g., threat of new entrants) might influence organization-level variables like organization strategy, which in turn might influence organization-level revenue. Instead of or in addition to this indirect effect, these industry-level variables might have direct interactive effects on organization-level revenue.

Whether such cross-level effects have no, partial, or full mediation by an intervening variable is an open question that probably depends on the specific revenue driver–revenue relation under consideration.

In contrast to treating industry as an interacting independent variable in a cross-level model, a researcher might take the approach of controlling for industry competition by treating industry as a control (additive dummy) variable and then testing for a bivariate relation between competition and revenue. This approach, however, is not valid because it does not provide analysis or evidence on how an industry-level variable influences an organization-level variable (Luft & Shields, 2003). Valid cross-level models also require that relevant theory at these various levels be used to develop a cross-level interaction prediction. Moreover, a valid test of a cross-level model requires that variable measurements and data analysis be consistent with the cross-level theoretical model. Similar cross-level issues occur for bottom-up models; for example, how customer- and/or product-level revenue drivers influence organization-level revenue.

6. CONCLUSION

This paper has reviewed revenue drivers based on five revenue-driver models in the accounting literature and located them in a model by reference to their levels of analysis. Analysis of this model indicates that most of the quantitative evidence on revenue drivers is at the organization level. We encourage future research to provide evidence on revenue drivers at the customer, product, and industry levels as they are likely to be important determinants of organization-level revenue. Moreover, in addition to the single-level models in the accounting literature, a complete understanding of revenue drivers is likely to require cross-level models.

Research could also be more informative if it provides analysis of and evidence on alternative relations between revenue drivers and revenue, beyond the modal relation examined to date (positive, linear, additive, direct, and contemporaneous). Related, no accounting study provides information on the duration of a revenue-driver effect or on whether revenue driver–revenue relations are bidirectional. Motivating and structuring research on revenue driver–revenue relations based on levels of analysis and seven other characteristics of relations (sign, linearity, additivity, directness, directionality, timing, and duration) as we have proposed can provide a broad and valid theoretical and empirical basis to investigate revenue drivers and, more generally, cost and profit drivers.

NOTES

1. The literature search used the databases Science Direct and Business Source Premier with the key words being revenue, revenue drivers, profit, and sales.
2. An interesting variable measure issue arises in this marketing research in that because revenue is product of unit sales price and sales quantity, using revenue as a variable has the potential to induce spurious correlation when, for example, sales price drives sales quantity (Hanssens et al., 2001). In this line of research, revenue is usually operationalized and measured as sales volume.

ACKNOWLEDGMENTS

The authors would like to thank two anonymous reviewers for their helpful comments for revising this paper and Alex Woods for his research assistance.

REFERENCES

Assmus, G., Farley, J. U., & Lehmann, D. R. (1984). How advertising affects sales: Meta-analysis of econometric results. *Journal of Marketing Research*, *21*(1), 65–74.

Babakus, E., Bienstock, C. C., & Van Scottter, J. R. (2004). Linking perceived quality and customer satisfaction to store traffic and revenue growth. *Decision Sciences*, *35*(4), 713–737.

Bain, J. S. (1968). *Industrial organization* (2nd ed.). New York: Wiley.

Banker, R., Chang, H., & Cunningham, R. (2003). The public accounting industry production function. *Journal of Accounting and Economics*, *35*(2), 255–281.

Banker, R., Konstans, C., & Mashruwala, R. (2000a). *A contextual study of links between employee satisfaction, employee turnover, customer satisfaction and financial performance*. Working Paper. University of Texas at Dallas.

Banker, R., & Mashruwala, R. (2001). *The importance of components of customer satisfactions in performance evaluation*. Working Paper. University of Texas at Dallas.

Banker, R., Potter, G., & Srinivasan, D. (2000b). An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review*, *75*(1), 65–92.

Barney, J. B. (1997). *Gaining and sustaining competitive advantage*. Reading, MA: Addison-Wesley.

Batt, R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. *Academy of Management Journal*, *45*(3), 587–597.

Behn, B., & Riley, R., Jr. (1999). Using nonfinancial information to predict financial performance: The case of the U.S. Airline industry. *Journal of Accounting, Auditing and Finance*, *14*(1), 29–56.

Blattberg, R. C., & Wisniewski, K. J. (1989). Price-induced patterns of competition. *Marketing Science*, *8*(4), 291–309.

Bronnenberg, B. J., Mahajan, V., & Vanhonacker, W. R. (2000). The emergence of market structure in new repeat-purchase categories: The interplay of market share and retailer distribution. *Journal of Marketing Research*, *37*(1), 16–31.

Bryant, L., Jones, D., & Widener, S. (2004). Managing value creation within the firm: An examination of multiple performance measures. *Journal of Management Accounting Research*, *16*, 107–131.

Bucklin, R. E., & Gupta, S. (1992). Brand choice, purchase incidence, and segmentation: An integrated modeling approach. *Journal of Marketing Research*, *29*(2), 201–215.

Cooper, R. (1990). Cost classification in unit-based and activity-based manufacturing cost systems. *Journal of Cost Management*, *4*(3), 4–14.

Dikolli, S., Kinney, W., & Sedatole, K. (2004). *Measuring trading partner relationships: Inferences from customer switching costs*. Working Paper, University of Texas at Austin.

Dikolli, S., & Sedatole, K. (2003). *The conditioning role of a customer satisfaction performance measure*. Working Paper. University of Texas at Austin.

Epstein, M., Kumar, P., & Westbrook, R. (2000). The drivers of customer and corporate profitability: Modeling, measuring, and managing the casual relationships. *Advances in Management Accounting*, *9*, 43–72.

Epstein, M., & Westbrook, R. (2001). Linking actions to profits in strategic decision making. *Sloan Management Review*, *42*(3), 39–49.

Foster, G., & Gupta, M. (1999). *The customer profitability implications of customer satisfaction*. Working Paper. Stanford University.

Foster, G., Gupta, M., & Sjoblom, L. (1996). Customer profitability analysis: Challenges and new directions. *Journal of Cost Management*, *10*(1), 5–17.

Hanssens, D., Parsons, L. J., & Schultz, L. (2001). *Market response models: Econometric and time series analysis* (2nd ed.). Boston, MA: Kluwer Academic Publishers.

Horngren, C., Datar, S., & Foster, G. (2006). *Cost accounting* (12th ed.). Upper Saddle River, NJ: Prentice-Hall.

Ittner, C., & Larcker, D. (1998a). Innovations in performance measurement: Trends and research implication. *Journal of Management Accounting Research*, *10*, 205–238.

Ittner, C., & Larcker, D. (1998b). Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research*, *36*(Supplement), 1–35.

Ittner, C., & Larcker, D. (2001). Assessing empirical research in management accounting: A value-based management perspective. *Journal of Accounting and Economics*, *32*, 349–410.

Ittner, C., Larcker, D., & Randall, T. (1997). The activity-based cost hierarchy production policies and firm profitability. *Journal of Management Accounting Research*, *9*, 143–162.

Ittner, C., Larcker, D., & Randall, T. (2003). Performance implications of strategic performance measurement in financial services firms. *Accounting, Organizations and Society*, *28*(7/8), 715–741.

Kaplan, R., & Cooper, R. (1998). *Cost and effect*. Boston, MA: Harvard Business School Press.

Kaplan, R., & Narayanan, V. (2001). Measuring and managing customer profitability. *Journal of Cost Management*, *15*(5), 5–15.

Kaplan, R., & Norton, D. (1996). *Translating strategy into action: The balanced scorecard*. Boston, MA: Harvard Business School Press.

Kaplan, R., & Norton, D. (2001). *The strategy-focused organization*. Boston, MA: Harvard Business School Press.

Kaplan, R., & Norton, D. (2004). *Strategy maps: Converting intangible assets into tangible outcomes*. Boston, MA: Harvard Business School Press.

Kennedy, P. A. (2003). *Guide to econometrics* (5th ed.). Cambridge, MA: MIT Press.

Koller, T. (1994). What is value-based management? *The McKinsey Quarterly*, *3*, 87–101.

Kopalle, P. K., Mela, C. F., & March, L. (1999). The dynamic effect of discounting on sales: Empirical analysis and normative pricing implications. *Marketing Science*, *18*(3), 317–332.

Leone, R. P. (1995). Generalizing what is known about temporal aggregation and advertising carryover. *Marketing Science*, *14*(3, Part 2 of 2), G141–G150.

Luft, J., & Shields, M. D. (2003). Mapping management accounting: Graphics and guidelines for theory-consistent empirical research. *Accounting, Organizations, and Society*, *28*(2/3), 169–249.

Martin, S. (2002). *Advanced industrial economics* (2nd ed.). Maldin, MA: Blackwell.

Mistry, J., & Johnston, H. (2004). *The impact of information technology in banking*. Working Paper. Babson College.

Nagar, V., & Rajan, M. (2001). The revenue implications of financial and operational measures of product quality. *The Accounting Review*, *76*(4), 495–513.

Narasimhan, R., & Kim, S. (2002). Effect of supply chain integration on the relationship between diversification and performance: Evidence from Japanese and Korean firms. *Journal of Operations Management*, *20*(3), 303–323.

Petersen, M. (2003). *Drivers of revenue in specialty outreach clinics*. Working Paper. Arizona State University.

Porter, M. (1980). *Competitive strategy*. New York: Free Press.

Porter, M. (1985). *Competitive advantage*. New York: Free Press.

Ramaswamy, V., Desarbo, W. E., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, *12*(1), 103–124.

Rangaswamy, A., Sinha, P., & Zoltners, A. (1990). An integrated model-based approach for sales force structuring. *Marketing Science*, *9*(4), 279–298.

Reibstein, D. J., & Ferris, P. W. (1995). Market share and distribution: A generalization, a speculation and some implications. *Marketing Science*, *14*(3, Part 2 of 2), G190–G202.

Rucci, A., Kirn, S., & Quinn, R. (1998). The employee-customer-profit chain at Sears. *Harvard Business Review*, *76*(1), 83–97.

Scherer, F., & Ross, D. (1990). *Industrial market structure and economic performance* (3rd ed.). Boston, MA: Houghton Mifflin Co.

Sethuraman, R., Srinivasan, V., & Kim, D. (1999). Asymmetric and neighborhood cross-price effects: Some empirical generalizations. *Marketing Science*, *18*(1), 23–41.

Sethuraman, R., & Tellis, G. J. (1991). An analysis of the tradeoff between advertising and price discounting. *Journal of Marketing Research*, *28*(2), 160–174.

Shank, J. N., & Govindarajan, V. (1989). *Strategic cost analysis*. Boston: Irwin.

Shank, J., & Govindarajan, V. (1992). Strategic cost management and the value chain. *Journal of Cost Management*, *5*(4), 5–21.

Shankar, V., Carpenter, G. S., & Krishnamurthi, L. (1999). The advantages of entry in the growth stage of the product life cycle: An empirical analysis. *Journal of Marketing Research*, *36*(2), 269–276.

Simon, H. (1982). Adpuls: An advertising model with wearout and pulsation. *Journal of Marketing Research*, *19*(3), 352–363.

Smith, R., & Wright, W. (2004). Determinants of customer loyalty and financial performance. *Journal of Management Accounting Research*, *16*, 183–205.

Tellis, G. (1988). The price elasticity of selective demand: A meta-analysis of econometric models of sales. *Journal of Marketing Research*, *25*(4), 331–341.

Trueman, B., Wong, M., & Zhang, X. (2001). Back to basics: Forecasting the revenues of internet firms. *Review of Accounting Studies*, *6*, 305–329.

FINANCIAL MEASURES BIAS IN THE USE OF PERFORMANCE MEASUREMENT SYSTEMS

Gerald K. DeBusk, Larry N. Killough and Robert M. Brown

ABSTRACT

This paper examines potential cognitive difficulties inherent in the use of performance measurement systems. We examine the potential for emphasizing financial measures as compared to nonfinancial measures in the evaluation of an organization's overall performance. The results suggest that users of performance measurement data will emphasize historical financial measures. Two separate experiments provide additional evidence that users of performance measurement data suffer a halo bias, in that an organization's performance on financial measures appears to influence their perception of the organization's performance on nonfinancial measures.

INTRODUCTION

Users of management accounting information have been critical of organizations whose performance measurement system is dominated by traditional financial measures. Traditional financial measures are often

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 61–89

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14003-9

characterized as being too historical or backward-looking (Ittner & Larcker, 1998b). In 1994, the American Institute of Certified Public Accountants (AICPA) called for organizations to report more forward-looking information and nonfinancial measures for key business processes. More recently, the American Accounting Association Financial Accounting Standards Committee (2002) urged the Financial Accounting Standards Board to encourage the reporting of nonfinancial measures in order to increase the relevance of external reporting.

Several studies have examined the value-relevance and predictive ability of nonfinancial performance information. Amir and Lev (1996) examined the value-relevance of nonfinancial information in the cellular phone industry. The nonfinancial information was positively associated with stock prices. Amir and Lev (1996, p. 5) state, "In the cellular industry, the value-relevance of nonfinancial information overwhelms that of traditional, financial indicators. …we expect this to be the case in other science-based, high-growth sectors." Ittner and Larcker (1998a) provide evidence that customer satisfaction is positively related to market value, and Dempsey, Gatti, Grinnell, and Cats-Baril (1997) provide evidence financial analysts use or want to use a wide range of nonfinancial information. These sentiments are echoed by Birchard (1995, p. 43) in his comments concerning a recent survey. He states, "According to a recent survey, 80 percent of large American companies want to change their performance measurement systems. No wonder. Yesterday's accounting results say nothing about the factors that actually help grow market share and profits – things like customer service innovation, R&D effectiveness, the percent of first-time quality, and employee development."

In order to overcome perceived limitations of managing solely with traditional financial measures, many firms have adopted new performance measurement systems that utilize a combination of financial and nonfinancial measures. Probably the most widely used of these new systems is the Balanced Scorecard (BSC), introduced by Kaplan and Norton in 1992. It combines financial and nonfinancial measures in a more "balanced" approach to performance measurement. A survey conducted by Renaissance Worldwide, Inc. estimated that 60% of Fortune 1000 companies have implemented or experimented with the BSC (Silk, 1998).

Management accountants have a role to play in the development, implementation, and use of these new performance systems. Barsky and Bremser (1999, p. 12) state, "Financial managers will be called upon to integrate diverse sets of data and provide sophisticated analysis and support for critical business decisions." Kaplan (1995) believes it is possible for

management accountants to become part of the management team, participating in the formulation and implementation of strategy. Management accountants can be involved in translating strategies into operational measures and the design of new information systems, according to Kaplan (1995). Additionally, Barsky and Bremser (1999) suggest management accountants should take the lead in the measurement and management of business risk. The Institute of Management Accountants' 1999 Practice Analysis, *Counting More*, *Counting Less*, provides evidence that management accountants are spending more time than ever before as internal consultants, business analysts, and even partners in decision-making processes. According to their survey, nearly 80% of management accountants spend more time actively involved in business decisions than 5 years earlier. Also, 82% of those surveyed expect to spend greater time in the next 3 years actively involved in business decisions (Siegel & Sorenson, 1999).

This paper examines cognitive difficulties and judgmental effects in the use of these types of performance measurement systems. We used accounting undergraduate and graduate students as a proxy for accounting professionals. Accountants are the chief custodian, user, and disseminator of performance measurement information and they, to a large extent, influence other managers' attitudes toward organizational performance. In this sense, accounting students are also a proxy for other professionals in the organization. We also used marketing students as subjects. These two groups, accounting and marketing students, combine to give the experiments a broad base of subjects and more generalizable findings.

Prior literature suggests that users of performance measurement systems, that combine financial and nonfinancial measures, experience cognitive difficulties when evaluating organizational performance (Lipe & Salterio, 2000; Ittner, Larcker, & Meyer, 2003; Banker, Chang, & Pizzini, 2004). The data presented in this paper suggest that users of performance measurement data emphasize the financial measures over the nonfinancial measures in evaluating overall organizational performance. Evidence is also provided of a halo effect, in that an organization's performance on financial measures appears to influence an individual's perception of the organization's performance on nonfinancial measures.

Accountants, it could be argued, are trained more in the use of financial measures than in nonfinancial measures and this may account for their bias toward financial measures. To answer this potential criticism, we used marketing students in an experiment to examine the potential halo bias. Evidence of a halo bias was found with the marketing students as well. The results of both experiments combine to provide strong evidence that users of performance

measurement data allow financial performance to influence their judgment concerning the organization's performance on nonfinancial measures.

The remainder of this paper is organized as follows. The following section contains a literature review and presents the research hypothesis, the next section discusses the methodology of the two experiments, and the section that follows presents the results. Following the results section is a section discussing implications of both experiments.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

To evaluate overall organizational performance, one must establish relative weights for each of the performance measures in a performance measurement system, a task requiring cognitive effort. Several studies provide evidence that users of performance measurement systems experience cognitive difficulties in evaluating organizational performance (e.g. Lipe & Salterio, 2000; Ittner et al., 2003; Banker et al., 2004). Subjects in the Lipe and Salterio (2000) study experienced a common measures bias. When evaluating two divisions within the same organization, the subjects relied on measures that were common to both divisions and failed to rely on measures unique to a particular division. Banker et al. (2004) confirmed the findings in Lipe and Salterio (2000) and found evidence on the importance of the linkage between strategy and the performance measures.[1]

Lipe and Salterio (2000) suggest that the bias for common measures may translate into a bias for financial measures since financial measures are often common between divisions. The weighting of financial versus nonfinancial measures is especially critical. Financial measures have been criticized as too historical and backward-looking. Nonfinancial measures are perceived to be predictive and forward-looking (Ittner & Larcker, 1998a, b) although sometimes perceived to suffer from poor measurement quality (Lingle & Schiemann, 1996). Nonfinancial measures have been found to be value relevant (Amir & Lev, 1996; Ittner & Larcker, 1998a). Recognizing these facts, the AICPA (1994) called for more extensive reporting of nonfinancial performance measures. Kaplan and Norton (2001, p. 376) proposed "nearly 80% of the measures on a Balanced Scorecard should be nonfinancial," suggesting that nonfinancial measures be emphasized in management decision-making and evaluations of performance.

While there is a trend toward a greater reporting of nonfinancial measures, there is some doubt as to whether nonfinancial measures receive the

proper weight in evaluating performance (Lipe & Salterio, 2000). "At Volvo, the principal challenge is considered to be that of giving nonfinancial measures as much weight as the more established financial ones," report Olve, Roy, and Wetter (1999, p. 119). Evidence of a financial measures bias was provided by a 1996 Towers Perrin survey that found BSC adopters placed an average of 56% of the relative weight on financial measures (Ittner & Larcker, 1998b). DeBusk, Brown, and Killough (2003) also found support for the finding that performance measurement system users view bottom-line financial measures as more important than nonfinancial measures.

Does the use of a set of measures overwhelm accountants and managers to such an extent that they must concentrate on traditional financial measures? Studies on information overload suggest that a large number of performance measures can hinder a manager's ability to evaluate the organization (Ittner & Larcker, 1998b). Some believe that managers may rely on what they know best, financial measures. Anthony and Govindarajan (2001, p. 451) state

> … not only are most senior managers well trained and adept with financial measures, they also keenly feel pressure regarding the financial performance of their companies. Shareholders are vocal, and boards of directors frequently apply pressure on the shareholders' behalf. This pressure may overwhelm the long-term, uncertain payback of the nonfinancial measures.

The anecdotal evidence (Olve et al., 1999), survey evidence (Ittner & Larcker, 1998b), and data from experiments (Lipe & Salterio, 2000; DeBusk et al., 2003) suggest that Govindarajan's analysis may be correct. Users of a performance measurement system may discount the importance of nonfinancial measures and rely more on financial measures. Therefore, the following research hypothesis is examined in this study.

> **H**. Financial measures will be emphasized, as compared to nonfinancial measures, in tasks involving the assessment of overall organizational performance.

One potential way for a financial measures bias to manifest itself is through a halo effect. If users of performance measurement data allow the organization's financial results to influence their perception of the organization's performance on nonfinancial measures, a halo effect or halo bias would exist. Pedhazur and Schmelkin (1991, p. 121) state, "One of the most common sources of bias is the halo effect, a constant error that occurs when raters' general impressions bias their ratings of distinct aspects of the rates." Evidence of a halo bias has been found in other studies involving performance measurement. In Nelson et al. (1992), it was noted that a halo effect caused hospital patients to rate the hospital as generally high or low on all measures. Herman and Renz (1999) expected to find multiple dimensions in measuring the effectiveness of

non-profit organizations. All of the measures in Herman and Renz (1999) loaded, however, on a single factor with a high degree of intercorrelation among the items (Cronbach alpha was 0.85). This unidimensionality and high intercorrelation are indicative of a halo effect. We examine, in this paper, potential halo biases as evidence of a financial measures bias.

METHOD

We conducted two experiments (experiments 1 and 2). The first experiment used accounting undergraduate and accounting graduate students to examine the potential bias toward financial measures including a potential halo bias. This experiment contained three different versions in a between-subjects design. Two of the versions were constructed in such a way that comparing the subject's perceptions of organizational performance between the versions would allow us to examine the question of emphasis on financial versus nonfinancial measures and to examine potential halo biases. The second experiment used marketing undergraduate students to also examine the potential halo bias. It contained two versions in a between-subjects design. The experimental manipulations were designed to facilitate the examination of a potential halo bias. The second experiment also had more subjects per cell than experiment 1, which facilitated the use of principal components analysis, a key method to examine halo effects.

Experiment 1

A case was developed giving sufficient background data, information on vision and strategies, and measures of performance against targets for a major brewery. The case was modified from an Institute of Management Accountants' case titled *Coors Case: Balanced Scorecard* by Hugh Grove, Tom Cook, and Ken Richter. Specific information was given on 24 individual performance measures including a definition of the measure, the target for the year, and the actual results for the year. The measures were organized onto one page and listed alphabetically.[2] Participants were given two packages of information. The first contained the background information (see Appendix A). The second contained the actual survey instrument and included the actual measures (see Appendix B). Participants in the experiment were asked to evaluate the organization's performance on individual measures and on an overall basis (ORGPERF) on a scale of 1 to 6, where one equals very poor and six equals very good. Participants, in

addition, were asked to assess the importance of each measure in evaluating the organization's performance on a scale of one to six, where one equals not important and six equals critical.

Three different measurement scenarios or versions (VER1–3) were utilized to alter performance against target.

Each version reflected in Table 1 was developed so the individual measures would support an overall scenario where return on invested capital (ROI or measure no. 16) was below target (VER1), on target (VER2), or above target (VER3). In addition to ROI, four other measures followed the same pattern where in VER1 actual was below target, in VER2 actual was on target, and in VER3 actual was better than target. The other measures were varied in an attempt to support the performance scenario in each version. Version 1 performance was generally below target, while version 2 performance was generally equal to target. Version 3 performance was better than target for net profit and ROI due to strength in selling, general and administrative costs but other measures were worse than target signaling future profitability problems.

The study was conducted with undergraduate and graduate students at a large state university in the eastern United States. Undergraduate students were obtained from two sections of a Cost Accounting course and graduate students from an Advanced Managerial Accounting course. The case and survey instruments were introduced after a 45–60 minute lecture on performance measurement systems focusing on the BSC system. The purpose of the lecture was to insure that the students were familiar with performance measurement systems like the BSC that emphasize nonfinancial measures. Students read the case and completed the surveys as part of a take-home exercise.

A total of 67 usable surveys were obtained. Ten surveys were dropped because of missing information. Ninety-one percent of the participants were accounting majors or had accounting listed as part of a double major. Of the 67 participants, 35 (52%) were female and 32 (48%) were male; 44 (66%) were undergraduate and 23 (34%) were graduate students. Graduate students comprise 43% of the version 1 sample, 29% of the version 2, and 30% of the version 3. Using MANOVA no significant differences in responses were noted between the two groups: undergraduate and graduate students.

Experiment 2

A possible limitation of the previous experiment is that it was conducted with accounting students whose training may predispose them to the use of financial measures. A second experiment using marketing students was undertaken to confirm the aforementioned results on halo effects. It was

Table 1. Organizational Performance Measures for Each Version – Experiment 1.

No.	Performance Measure	Target	Actual 2000 Version			Actual Versus Target[a] Version		
			1	2	3	1	2	3
1	Annual market share increase	0.05%	0.04%	0.05%	0.04%	−20%	0%	−20%
2	Barrels produced per DLH	6.0	5.0	6.0	5.8	−17%	0%	−3%
3	Baseline growth (for key brands and markets)	2.0%	2.0%	2.0%	2.0%	0%	0%	0%
4	Beer waste & package scrap	0%	0.5%	0.5%	0.5%			
5	Community involvement (volunteer hours per employee annually)	30	25	30	25	−17%	0%	−17%
6	Customer complaints (per 100,000 barrels sold)	0.05%	0.06%	0.05%	0.04%	−20%	0%	20%
7	Incremental growth (for high potential brands and markets)	3.0%	2.5%	3.0%	2.5%	−17%	0%	−17%
8	Load item accuracy	100%	93%	95%	96%	−7%	−5%	−4%
9	Load schedule performance	100%	55%	65%	60%	−45%	−35%	−40%
10	Manufacturing cost per barrel	$53	$55	$53	$54	−4%	0%	−2%
11	Net profit per barrel	$6	$3	$6	$8	−50%	0%	33%
12	New products (new brands introduced each year)	6	5	6	5	−17%	0%	−17%
13	Plant productivity	80%	76%	80%	78%	−5%	0%	−3%
14	Production stability	100%	50%	65%	60%	−50%	−35%	−40%
15	Quality index (out of 100)	90	85	90	95	−6%	0%	6%
16	Return on invested capital	12.0%	7.1%	12.0%	14.5%	−41%	0%	21%
17	Revenue per barrel	$100	$100	$100	$100	0%	0%	0%
18	Safety (lost work incident rate)	0	0.0001	0.0001	0.0001			
19	Safety (total case incident rate)	0	0.0002	0.0002	0.0002			
20	SG&A cost per barrel	$27	$28	$27	$24	−4%	0%	11%
21	Skills (inventory of cross-functional employee skills)	7	6	7	6	−14%	0%	−14%
22	Throughput per month (millions of barrels)	1.75	1.72	1.75	1.74	−2%	0%	−1%
23	Training (hours per employee annually)	40	42	40	38	5%	0%	−5%
24	Warehouse moves (actual shipments as a percentage of plan)	100%	95%	95%	95%	−5%	−5%	−5%
	Average					−16%	−4%	−5%
	Average of nonfinancial measures					−15%	−5%	−10%
	Average of financial measures					−20%	0%	13%

[a]For measures 4, 18, and 19, the target is 0, which presents a problem in the computation of the percentage difference (division by 0). Averages ignore these three variables.

believed that marketing students would not possess the financial measures bias that possibly afflicts accounting students.

A total of 78 marketing students at a large state university in the eastern United States were given the survey instrument contained in Appendix C. Three surveys were discarded due to failure to follow instructions or missing data, leaving 75 usable surveys. Of the 75 participants, 53 (71%) were juniors and 22 (29%) were seniors. The survey was conducted in class and required approximately 15–20 min. Subjects received extra class credit and a commemorative postage stamp for their efforts.

Participants were given a case (see Appendix C), which had two versions (VER1 and VER2). Between the two versions, nonfinancial measures were unchanged and were close to targeted performance, while financial measures were on average 16.1% below target in VER1 and 16.1% above target in VER2, as portrayed in Table 2.[3] The study participants were asked to assess the organization's overall performance and performance on 20 individual measures on a 1–10 scale.

RESULTS

Experiment 1: Manipulation Checks

Table 3 provides descriptive statistics for each of the 24 dependent variables. There were no differences between versions in the importance of individual measures.[4]

As a validation that the different versions did portray differences in performance, a MANOVA was run on the participant's evaluation of the organization's performance on the 24 individual measures. The results were significant ($p = 0.000$) indicating perceived differences in performance among the three versions' 24 measures. The participants could differentiate performance among the versions but there was no differentiation in importance of the various 24 measures among the versions. Table 4 shows the mean evaluation scores on the individual performance measures[5] and on the organization's overall performance.[6]

Experiment 1: Data Analysis

Evidence concerning the research hypothesis was obtained by examining performance evaluations between versions 2 and 3 in experiment 1. Versions 2 and 3 were selected for analysis because of the tension provided by comparing

Table 2. Organizational Performance Measures for each Version – Experiment 2.

Performance Measure	Target	Actual Year 2003		Actual Versus Target	
		VER1	VER2	VER1	VER2
Average lead time in days (order to delivery)	7	7.2	7.2	−2.9%	−2.9%
Customer satisfaction index (out of 100)	90	91	91	1.1%	1.1%
Defects per 1,000 units	50	49.5	49.5	1.0%	1.0%
Employee satisfaction index (out of 100)	90	91	91	1.1%	1.1%
Employee turnover	2%	1.9%	1.9%	5.0%	5.0%
Free cash flow ($000)	$5,000	$3,989	$6,011	−20.2%	20.2%
Inventory turns	20	20	20	0.0%	0.0%
Labor efficiency	90%	89.5%	89.5%	−0.6%	−0.6%
Manufacturing cycle time in days	4	4.1	4.1	−2.5%	−2.5%
Market share	20%	20.2%	20.2%	1.0%	1.0%
Number of warranty claims per 1,000 shipments	40	39.7	39.7	0.8%	0.8%
On-time delivery	95%	95.5%	95.5%	0.5%	0.5%
Operating income ($000)	$10,000	$7,993	$12,007	−20.1%	20.1%
Percentage involvement in voluntary quality circles	80%	81.6%	81.6%	2.0%	2.0%
Purchase price variance – favorable (unfavorable)	1.0%	1.02%	1.02%	2.0%	2.0%
Return on investment	12%	9.6%	14.4%	−20.0%	20.0%
Revenue growth rate	10%	8.0%	12.0%	−20.0%	20.0%
Revenue per unit	$1,000	$1,000	$1,000	0.0%	0.0%
Safety (OSHA reportable lost time accidents)	2	2	2	0.0%	0.0%
Training hours per employee	40	39	39	−2.5%	−2.5%
Average				−3.7%	4.3%
Average of nonfinancial measures				0.4%	0.4%
Average of financial measures				−16.1%	16.1%

Table 3. Descriptive Statistics – Perceived Importance of Individual Measures for Experiment 1 ($n = 23$, 21, and 23 for Versions 1, 2, and 3, respectively).

No.	Performance Measure	Mean Importance Version				Standard Deviation Version			
		1	2	3	All	1	2	3	All
1	Annual market share increase	4.83	4.52	5.17	4.85	0.887	1.327	0.717	1.019
2	Barrels produced per DLH	4.57	4.29	4.43	4.43	0.728	1.231	1.121	1.033
3	Baseline growth (for key brands and markets)	4.57	4.81	4.70	4.69	1.080	0.981	0.926	0.988
4	Beer waste & package scrap	3.78	3.76	4.04	3.87	1.126	1.136	1.261	1.166
5	Community involvement (volunteer hours per employee annually)	2.74	2.95	2.96	2.88	1.176	1.396	1.065	1.200
6	Customer complaints (per 100,000 barrels sold)	5.00	4.81	5.13	4.99	0.853	1.078	0.757	0.896
7	Incremental growth (for high potential brands and markets)	4.57	4.71	4.96	4.75	0.945	1.007	0.767	0.910
8	Load item accuracy	4.48	4.10	4.39	4.33	0.898	0.889	0.891	0.894
9	Load schedule performance	4.26	4.10	4.09	4.15	0.964	0.889	0.900	0.909
10	Manufacturing cost per barrel	5.17	4.95	5.04	5.06	0.778	0.805	0.706	0.756
11	Net profit per barrel	5.22	5.05	5.43	5.24	0.671	0.805	0.662	0.720
12	New products (new brands introduced each year)	4.13	4.14	4.22	4.16	1.180	1.195	0.998	1.109
13	Plant productivity	4.65	4.86	4.70	4.73	0.775	0.964	0.822	0.845
14	Production stability	4.57	4.48	4.43	4.49	0.843	0.873	0.728	0.805
15	Quality index (out of 100)	4.96	4.71	4.52	4.73	0.928	0.902	1.238	1.038
16	Return on invested capital	4.57	5.00	5.39	4.99	1.273	1.049	0.722	1.080
17	Revenue per barrel	4.87	4.90	4.96	4.91	0.968	0.889	1.107	0.981
18	Safety (lost work incident rate)	4.04	4.48	4.09	4.19	1.065	1.123	0.949	1.048
19	Safety (total case incident rate)	4.22	4.10	3.91	4.07	1.043	1.338	1.083	1.146
20	SG&A cost per barrel	4.61	4.33	4.52	4.49	0.891	1.017	1.039	0.975
21	Skills (inventory of cross-functional employee skills)	4.57	4.67	4.48	4.57	0.945	1.017	0.947	0.957
22	Throughput per month (millions of barrels)	4.22	4.76	4.48	4.48	0.850	0.889	0.994	0.927
23	Training (hours per employee annually)	4.65	4.81	4.35	4.60	1.112	0.981	0.982	1.031
24	Warehouse moves (actual shipments as a percentage of plan)	4.13	4.24	3.91	4.09	1.254	0.831	1.379	1.177
	Average	4.47	4.48	4.51	4.49				
	Average of nonfinancial measures	4.37	4.38	4.37	4.37				
	Average of financial measures	4.89	4.85	5.07	4.94				

Table 4. Descriptive Statistics – Perceived Performance on Individual Measures and Overall for Experiment 1 ($n = 22$, 21, and 22 for Versions 1, 2, and 3, respectively).

No.	Performance Measure	Mean Evaluation Version				Standard Deviation Version			
		1	2	3	All	1	2	3	All
1	Annual market share increase	4.27	4.76	4.73	4.58	1.120	0.995	0.767	0.983
2	Barrels produced per DLH	3.91	4.48	4.68	4.35	0.684	1.030	0.839	0.909
3	Baseline growth (for key brands and markets)	4.86	4.76	5.36	5.00	0.990	0.889	0.727	0.901
4	Beer waste & package scrap	3.86	3.76	3.68	3.77	1.037	1.338	1.287	1.209
5	Community involvement (volunteer hours per employee annually)	4.05	4.57	3.95	4.18	0.844	1.165	0.899	0.998
6	Customer complaints (per 100,000 barrels sold)	4.27	4.62	5.27	4.72	1.202	1.161	0.767	1.125
7	Incremental growth (for high potential brands and markets)	3.73	4.67	4.32	4.23	0.935	1.017	0.894	1.012
8	Load item accuracy	3.82	3.95	4.27	4.02	1.220	0.921	0.767	0.992
9	Load schedule performance	2.23	2.52	2.77	2.51	1.572	1.365	1.152	1.371
10	Manufacturing cost per barrel	4.05	4.67	4.59	4.43	0.844	0.913	1.008	0.951
11	Net profit per barrel	2.86	4.90	5.50	4.42	1.167	1.044	0.859	1.530
12	New products (new brands introduced each year)	4.50	4.48	4.64	4.54	0.740	1.327	0.658	0.937
13	Plant productivity	4.14	4.52	4.68	4.45	1.037	0.928	0.716	0.919
14	Production stability	2.18	2.76	2.95	2.63	1.181	1.375	1.290	1.306
15	Quality index (out of 100)	3.95	4.62	5.23	4.60	0.722	0.865	0.922	0.981
16	Return on invested capital	3.18	4.90	5.50	4.52	0.795	0.995	0.673	1.288
17	Revenue per barrel	5.23	5.05	5.05	5.11	0.869	0.921	0.950	0.904
18	Safety (lost work incident rate)	5.09	4.38	4.64	4.71	0.750	0.805	1.049	0.914
19	Safety (total case incident rate)	5.00	4.29	4.73	4.68	0.816	0.902	0.935	0.920
20	SG&A cost per barrel	4.36	4.57	5.09	4.68	1.002	1.207	1.065	1.120
21	Skills (inventory of cross-functional employee skills)	4.09	4.67	4.45	4.40	0.811	0.913	0.739	0.844
22	Throughput per month (millions of barrels)	4.32	4.86	4.86	4.68	0.894	0.910	0.710	0.868
23	Training (hours per employee annually)	5.27	4.95	4.68	4.97	0.767	0.740	0.780	0.790
24	Warehouse moves (actual shipments as a percentage of plan)	4.23	4.19	4.41	4.28	0.813	0.750	0.854	0.801
	Overall organization performance	4.35	4.52	4.78	4.55	0.647	0.680	0.422	0.610
	Average	4.07	4.42	4.59	4.36				
	Average of nonfinancial measures	4.09	4.31	4.44	4.28				
	Average of financial measures	3.94	4.82	5.15	4.63				

the financial and nonfinancial measures between the two versions. Versions 2 and 3 each outperform version 1 on both financial and nonfinancial measures. However, a comparison of versions 2 and 3 is a study where each version outperforms the other on one category (financial or nonfinancial) of measures. The average performance for the nonfinancial measures is 5% below target in version 2 but 10% below target in version 3; however, financial measures were on target in version 2 (0%) but 13% better than target in version 3.[7]

In version 2, all of the financial measures and most of the nonfinancial measures indicate performance exactly equal to target. In version 3, the favorable financial measures of net profit and ROI are driven by selling, general & administrative (SG&A) costs being $3 (11%) better than target. Offsetting some of the SG&A savings is worse than targeted results in manufacturing cost. Barrels produced per direct labor hour, throughput per month, load schedule performance, production stability, and plant productivity are all worse in version 3 than in version 2, indicating poor manufacturing and overall productivity in version 3 compared to target and version 2. Worse than targeted and version 2 performances in training and skills measures might help explain some of the productivity problems in version 3 and possibly indicate yet additional future declines in productivity. Cuts in training might have accounted for some of the version 3 savings on the SG&A line, but could lead to more manufacturing problems in future years. Cuts in advertising might also account for some of the lower SG&A costs in version 3 and some of the growth problems. Version 3 is also behind version 2 and below target in annual market share increase, incremental growth for high potential brands and markets, and new products introduced each year. These results suggest problems for the future in maintaining revenue growth. Version 3 outperforms version 2 on only three nonfinancial measures: customer complaints (0.04% VER3; 0.05% VER2), quality index (95% VER3; 90% VER2), and load item accuracy (96% VER3; 95% VER2).

The nonfinancial measures reflected in Table 1 indicate that version 2 is superior to version 3 (version 2 averages 5% worse than target while version 3 averages 10% worse than target). The poor performance in version 3 within the nonfinancial measures signals a future decline in financial performance even though the VER3 financial measures reflected in Table 1 are better than those in version 2 in the current period (average of 13% favorable to target in VER3 versus on target in VER2).

Average performance of all measures is 4% worse than target in version 2 and 5% worse than target in version 3. If each measure is weighted equally, version 2 should be perceived as outperforming version 3. Considering the recommendations for a "balanced approach" and that nonfinancial

measures are said to predict future financial performance, one could argue that the two versions should be perceived as roughly equivalent or version 2 as slightly superior.

It is difficult to determine the relative weights of individual measures subjects use in making an evaluation of organizational performance; however, we may find evidence of weighting by examining their assessment of overall organizational performance. If the study participants assess version 3 as superior to version 2, it would suggest that nonfinancial measures were not emphasized. We can analyze the perception of the performance of the two versions by examining responses on the overall organizational performance variable (ORGPERF). A *t*-test was performed on ORGPERF (version 2 compared to version 3) to test for a difference in perceived overall performance.

Experiment 1 – Results

The results of the *t*-tests to determine if subjects perceive a difference in the overall organizational performance between versions 2 and 3 are given in Table 5. The research participants receiving version 3 evaluated the organization better than those receiving version 2 ($p = 0.037$, one-tailed).[8] These results support the hypothesis and suggest an emphasis on financial measures and a lack of emphasis on nonfinancial measures.

A paired *t*-test was performed comparing the mean importance of the five financial measures (4.94) to the mean importance of the 19 nonfinancial variables (4.37). The mean importance of the financial measures was significantly higher ($p = 0.000$).

Experiment 1 – Evidence of Halo Effects

Further analysis reveals that an emphasis on financial results was not the only item driving the view that version 3 outperformed version 2. The results

Table 5. Version 2 Versus Version 3 Overall Performance Assessment *t*-Test for Experiment 1.

	Mean	Standard Deviation	*N*	Sig. (1-tailed)
VER2	0.52	0.512	21	
VER3	0.78	0.422	23	
Difference	−0.26			0.037[a]

Note: A binary split was done on overall organizational performance (0 if ORGPERF $\leqslant 4$; 1 if ORGPERF $\geqslant 5$).

[a]Nonparametric tests yielded similar results.

in Table 4 reveal that participants view version 3 as outperforming version 2 on the nonfinancial measures as well (mean of 4.44 versus 4.31). This is intriguing considering the average performance to target, shown in Table 1, is worse in version 3 than version 2 (VER3 -10%; VER2 -5%). See Fig. 1 for an illustration of the findings. The perceived performance on the nonfinancial measures appears to be more related to the financial measures than the nonfinancial measures.

A MANOVA performed on the assessed performance on the 19 nonfinancial measures indicates a significant difference ($p = 0.001$) between participants' assessment of performance on the two versions.[9] Fig. 1 can be explained by version 3 subjects forming a general impression of the organization based on the financial results. This general impression biases their ratings on the version 3 nonfinancial results, a halo effect or halo bias.

There are too few subjects in each version to perform a factor analysis. However, Cronbach alphas for each version ranged from 0.85 to 0.90. In version 3, one should not expect perception of financial performance to be correlated with perceived nonfinancial performance since financial performance was above target on average and nonfinancial performance was below target on average. The subject's overall perception of financial performance in version 3 was, however, highly correlated with their perception of performance on the nonfinancial measures (Pearson correlation equals 0.42; $p = 0.052$, two-tailed) and thus provides more evidence of a halo bias.

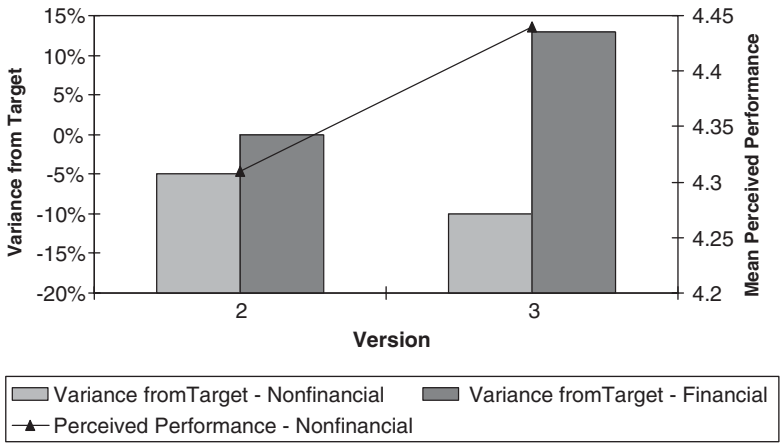


Fig. 1. Mean Perceived Performance on Nonfinancial Measures – Experiment 1 (Compared to Variance from Target for both Financial and Nonfinancial Measures).

Experiment 2 – Results

The second experiment used marketing undergraduate students as subjects. It was believed that marketing students would not possess the financial measures bias that possibly afflicts accounting students. Like experiment 1, the marketing students were asked to assess the organization's overall performance and their performance on several individual measures. The results are presented in Table 6.

A perceived difference by the participants in the performance on the nonfinancial measures would be an evidence of a halo effect. The average perceived performance was 7.66 for VER1 and 8.08 for VER2. A *t*-test was performed on the perceived performance data indicating a significant

Table 6. Descriptive Statistics – Perceived Performance on Individual Measures and Overall/Test of Halo Effects ($n = 38$ for VER1, $n = 37$ for VER2).

Performance Measure	Mean Evaluation			Standard Deviation		
	VER1	VER2	All	VER1	VER2	All
Average lead time in days (order to delivery)	7.66	7.59	7.63	1.56	1.82	1.68
Customer satisfaction index (out of 100)	8.26	8.68	8.47	1.13	1.03	1.09
Defects per 1,000 units	7.42	8.03	7.72	2.02	1.74	1.90
Employee satisfaction index (out of 100)	8.32	8.51	8.41	1.25	1.10	1.18
Employee turnover	7.45	8.32	7.88	1.91	1.33	1.70
Free cash flow ($000)	5.24	8.68	6.93	2.50	1.76	2.76
Inventory turns	7.79	8.27	8.03	1.74	1.54	1.65
Labor efficiency	7.13	7.35	7.24	1.98	1.84	1.90
Manufacturing cycle time in days	6.84	7.43	7.13	2.02	1.79	1.92
Market share	8.32	8.43	8.37	1.38	1.21	1.29
Number of warranty claims per 1,000 shipments	7.42	7.78	7.60	1.67	1.55	1.61
On-time delivery	8.29	8.46	8.37	1.59	1.22	1.41
Operating income ($000)	5.16	8.70	6.91	2.14	1.90	2.69
Percentage involvement in voluntary quality circles	7.74	8.24	7.99	1.64	1.32	1.50
Purchase price variance – favorable (unfavorable)	7.21	7.89	7.55	1.86	1.37	1.66
Return on investment	5.03	8.97	6.97	2.12	1.42	2.68
Revenue growth rate	5.39	8.86	7.11	2.16	1.40	2.52
Revenue per unit	7.74	8.16	7.95	1.61	1.50	1.56
Safety (OSHA reportable lost time accidents)	7.63	8.03	7.83	1.68	1.59	1.64
Training hours per employee	7.50	8.11	7.80	1.83	1.35	1.63
Overall organizational performance	7.47	8.57	8.01	1.48	1.12	1.42
Average	7.19	8.24	7.71			
Average of nonfinancial measures	7.66	8.08	7.87			
Average of financial measures	5.71	8.68	7.17			

difference ($p = 0.065$, one-tailed).[10] Further evidence is provided in Table 6 by examining the perceived performance on each of the 15 nonfinancial measures. All but one (average lead time) of the nonfinancial measures have a lower mean in VER1 than in VER2. The correlation between financial and nonfinancial measures provides further evidence of a halo effect by suggesting that the financial measures influence the perception of performance on the nonfinancial measures. Financial measures and nonfinancial measures had a correlation coefficient of 0.65 ($p = 0.000$).[11]

Principal components analyses were performed on the individual perceived performance variables in both versions. The results of the two principal components analyses are presented in Table 7. The factor loadings and Cronbach alphas presented in Table 7 provide strong evidence of unidimensionality and a high degree of inter-correlation, characteristics of a halo bias.

Table 7. Principal Components Analyses by Version – Experiment 2 ($n = 38$ for VER1, $n = 37$ for VER2).

Performance Measure	Factor Loadings	
	VER1	VER2
Average lead time in days (order to delivery)	0.478	0.722
Customer satisfaction index (out of 100)	0.600	0.708
Defects per 1,000 units	0.663	0.651
Employee satisfaction index (out of 100)	0.804	0.655
Employee turnover	0.675	0.716
Free cash flow ($000)	0.778	0.570
Inventory turns	0.782	0.763
Labor efficiency	0.784	0.800
Manufacturing cycle time in days	0.680	0.778
Market share	0.637	0.829
Number of warranty claims per 1,000 shipments	0.676	0.776
On-time delivery	0.720	0.762
Operating income ($000)	0.766	0.496
Percentage involvement in voluntary quality circles	0.604	0.871
Purchase price variance – favorable (unfavorable)	0.759	0.790
Return on investment	0.766	0.684
Revenue growth rate	0.779	0.678
Revenue per unit	0.858	0.807
Safety (OSHA reportable lost time accidents)	0.818	0.756
Training hours per employee	0.670	0.874
Cronbach alpha	0.948	0.951

DISCUSSION AND CONCLUSION

The evidence presented in this paper supports prior literature in suggesting that users of performance measurement systems may emphasize financial measures. The forward-looking nature of nonfinancial measures has been a major factor in the development of performance measurement systems that contain both financial and nonfinancial measurements. We found that subjects failed to emphasize the nonfinancial measures. We further found that the subjects allowed the financial results to bias their view of the nonfinancial results (i.e. a halo bias). The results suggest that when confronted with multiple measures of organizational performance, a person will form a general impression based on the financial results. This general impression will distort their perception of performance on the remaining nonfinancial measures. This bias was present even in the presence of a potential debiaser in the first study; participants were given training that promoted the predictive ability of nonfinancial measures.

This paper conducted two separate experiments using students with widely differing backgrounds (accounting students versus marketing students). The similar results found using two different subject pools enhances the generalizability of our findings. However, the generalizability of our findings may be hampered because students were used as subjects. Students lack the experience of managers and may use different criteria in evaluating an organization. Also, they may be more susceptible to the effects of information overload due to their lack of experience. Future research could extend these findings using experienced managers.

NOTES

1. Libby, Salterio, and Webb (2003) and Roberts, Albright, and Hibbets (2004) examined potential debiasing techniques in experiments similar to Lipe and Salterio (2000) with favorable results.

2. This study does not intend to test the effects of organizing measures into the BSC format. Interested parties should consult Lipe and Salterio (2002). In their study, BSC format had no effect in evaluations of two division managers when differences were scattered across the four BSC categories. However, Lipe and Salterio (2002) found that BSC format moderated the evaluations when the differences were concentrated into one BSC category.

3. The financial measures were free cash flow, operating income, ROI, revenue growth rate, and revenue per unit.

4. A MANOVA was performed to test if there were differences among the three versions in the relative importance of the measures. The significance levels for the test

statistics range from 0.339 to 0.586 indicating that no significant differences were detected on the relative importance of the individual measures among the three versions of the case. Observed power levels using alpha = 0.05 were above the recommended 0.80 threshold (Hair, Anderson, Tatham, & Black, 1998) for all test statistics but Roy's Largest Root, which had an observed power of 0.71.

5. Two of the 67 cases had missing values on the performance variables. Thus Table 4 reflects the means of a total of 65 cases (22 VER1, 21 VER2, and 22 VER3).

6. Utilizing ANOVA to examine the overall performance variable (ORGPERF), further evidence is provided of significant differences in the versions ($p = 0.050$). The overall performance scores and the average of the 24 individual measures demonstrate that the participants view performance in VER1, VER2, and VER3 differently.

7. See Table 1. For measures 4, 18, and 19, the target is 0 which presents a problem in the computation of the percentage difference (division by 0). However, actual performance is the same in all three versions. Averages ignore these three variables. Financial variables are 10, 11, 16, 17, and 20.

8. Because 42 of the 44 participants responded with either a 4 or 5 for the rating of organizational performance, splitting the data into two groups ($\leqslant 4$ and $\geqslant 5$) and performing the *t*-test is appropriate. Nonparametric Mann–Whitney tests yielded similar results.

9. Version 3 did outperform version 2 on three nonfinancial measures. To determine if this was causing the effect, another analysis was conducted excluding those nonfinancial measures where version 3 outperformed version 2. On the 16 nonfinancial measures where version 3 performed the same as or worse than version 2, the mean performance was 4.29 for version 2 and 4.35 for version 3. MANOVA results indicated a significant difference ($p = 0.005$).

10. Non-parametric tests yielded similar results.

11. If revenue per unit is excluded since it was unchanged between the versions, the correlation coefficient would be 0.57 ($p = 0.000$). Nonparametric correlations yield similar results.

ACKNOWLEDGMENTS

The authors wish to acknowledge the assistance of Thomas Klammer, William Kerler III, and the attendees at the 2003 AAA Southeast Regional Conference. In addition, we acknowledge the helpful comments of John Lee and two anonymous reviewers.

REFERENCES

American Accounting Association (2002). Recommendations on disclosure of nonfinancial performance measures. *Accounting Horizons*, *16*(4), 353–362.

American Institute of Certified Public Accountants (AICPA) (1994). *Improving business reporting – a customer focus*. New York, NY: AICPA.

Amir, E., & Lev, B. (1996). Value-relevance of nonfinancial information: The wireless communications industry. *Journal of Accounting & Economics*, *22*, 3–30.

Anthony, R. N., & Govindarajan, V. (2001). *Management control systems* (10th ed). New York, NY: McGraw-Hill.

Banker, R. D., Chang, H., & Pizzini, M. J. (2004). The balanced scorecard: Judgmental effects of performance measures linked to strategy. *The Accounting Review*, *79*(1), 1–23.

Barsky, N. P., & Bremser, W. G. (1999). Performance measurement, budgeting and strategic implementation in the multinational enterprise. *Managerial Finance*, *25*(2), 3–15.

Birchard, B. (1995). Making it count. *CFO*, *11*(10), 42–51.

DeBusk, G. K., Brown, R. M., & Killough, L. N. (2003). Components and relative weights in utilization of performance measurement systems like the balanced scorecard. *British Accounting Review*, *35*(3), 215–231.

Dempsey, S. J., Gatti, J. F., Grinnell, D. J., & Cats-Baril, W. L. (1997). The use of strategic performance variables as leading indicators in financial analysts' forecasts. *The Journal of Financial Statement Analysis*, *2*(4), 61–79.

Grove, H., Cook, T., & Richter, K. (2000). Coors case: Balanced scorecard. In: W. Bremser & J. Mackey (Eds), *Cases from management accounting practice*, Vol. 15. Montvale, NJ: Institute of Management Accountants.

Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice-Hall.

Herman, R. D., & Renz, D. O. (1999). Theses on nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, *28*(2), 107–126.

Ittner, C. D., & Larcker, D. F. (1998a). Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research*, *36*(Suppl), 1–35.

Ittner, C. D., & Larcker, D. F. (1998b). Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research*, *10*, 205–238.

Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review*, *78*(3), 725–758.

Kaplan, R. S. (1995). New roles for management accountants. *Journal of Cost Management*, *9*(3), 6–13.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard – measures that drive performance. *Harvard Business Review*, *70*(1), 71–79.

Kaplan, R. S., & Norton, D. P. (2001). *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Boston, MA: Harvard Business School Publishing Corporation.

Libby, T., Salterio, S. E., & Webb, A. (2003). The balanced scorecard: The effects of assurance and process accountability on managerial judgment. Working paper, Wilfrid Laurier University.

Lingle, J. H., & Schiemann, W. A. (1996). From balanced scorecard to strategic gauges: Is measurement worth it? *Management Review*, *85*(3), 56–61.

Lipe, M. S., & Salterio, S. E. (2000). The balanced scorecard: The judgmental effects of common and unique performance measures. *The Accounting Review*, *75*(3), 283–298.

Lipe, M. S., & Salterio, S. E. (2002). A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society*, *27*, 531–540.

Nelson, E. C., Rust, R. T., Zahorik, A., Rose, R. L., Batalden, P., & Siemanski, B. A. (1992). Do patient perceptions of quality relate to hospital performance? *Marketing Health Services*, *12*(4), 6–13.

Olve, N., Roy, J., & Wetter, M. (1999). *Performance drivers: A practical guide to implementing the balanced scorecard*. West Sussex, England: Wiley.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Roberts, M. L., Albright, T. L., & Hibbets, A. R. (2004). Debiasing balanced scorecard evaluations. *Behavioral Research in Accounting*, *16*, 75–88.

Siegel, G., & Sorenson, J. E. (1999). *Counting More, Counting Less: Transformations in the management accounting profession*. Montvale, NJ: Institute of Management Accountants.

Silk, S. (1998). Automating the balanced scorecard. *Management Accounting*, *79*(11), 38–44.

APPENDIX A. SELECTED CASE MATERIALS FOR EXPERIMENT 1

Company Background (Adapted from Coors Case: Balanced Scorecard; Grove, Cook, & Richter, 2000.)

Coors Brewing Company

Company Background. Coors had been a family owned and operated business from its inception in 1873 until 1993 when the first non-family member became President and Chief Operating Officer. However, Coors family members still held the Chairman of the Board of Directors and Chief Executive Officer positions and all voting stocks. Only non-voting, Class B common stock was publicly traded. Coors has been financed primarily by equity and has only borrowed capital twice in its corporate history. The first long-term debt, $220 million, 8.5% notes, was issued in 1991 and the final $40 million of principal was repaid by the end of 1999. The second long-term debt, $100 million, 7% unsecured notes, was issued in a 1995 private placement. $80 million of this principal is due in 2002 and the last $20 million is due in 2005.

In the mid-1970s, Coors was a regional brewery with an 11-state market, selling one brand in a limited number of packages through approximately 200 distributors. Traditionally, Coors beer had been a non-pasteurized, premium beer (however, with a recently developed sterilization process, its products now have the same shelf life as its competitors' pasteurized products). Coors plant in Golden, Colorado was its only production facility and it had no other distribution centers.

Over the next 25 years, Coors changed dramatically by expanding into all 50 states and various foreign markets. By the end of the twentieth century, Coors had production facilities in Golden, Colorado, Memphis, Tennessee, Elkton, Virginia, and Zaragoza, Spain. It had expanded to using 21 "satellite redistribution centers" in the United States before a special project reduced this number to eight. Beer shipments were made by both truck and railroad cars. Coors had approximately 650 domestic beer distributors although about 200 of them accounted for 80% of Coors total sales. Coors also had several joint ventures and international distributors in Canada, the Caribbean, Latin American, Europe, and the Pacific.

Coors had 16 beer brands, including a specialty line, Blue Moon that competed with the domestic micro brewing industry. However, Coors continued to focus upon its four key premium brands, Coors Light, Original Coors, Killian's Irish Red, and Zima. Coors Light was the fourth largest selling beer in the U.S. In packaging, Coors had to compete with the major competitors' value packaging, such as 12-packs and 30-packs. In 1959, Coors introduced the nation's first all-aluminum beverage can and in the late 1990s, it had introduced a baseball bat bottle and a football pigskin bottle. There were also numerous state-labeling laws to meet, such as returnable information, and packaging graphics to reinforce the Rocky Mountains image for Coors beer.

Competition in the beer industry was strong, especially in the United States. Anheuser–Busch (A/B) was the market leader with approximately 44% of the U.S. market, 80 million barrels sold, $8 billion beer sales, and $1 billion net profit. Due to its size, A/B was the acknowledged price leader in the industry. A/B also had 13 domestic production plants, including one in Ft. Collins, Colorado, to achieve its customer service goal of having no major domestic distributor more than 500 miles away from one of its beer production plants.

Number two in this market was Miller, owned by Philip Morris, with approximately 22% market share, 40 million barrels sold, $4 billion beer sales, and $460 million net profit. Miller also had seven domestic production plants. Coors was number three with an 11% market share, 20 million barrels sold, $2 billion beer sales, and $80 million net profit. Coors had three production plants in the United States. Its Colorado plant was the largest brewery in the world and served 70% of the U.S. market with its ten can lines, six bottle lines, and two keg lines.

There were no other domestic brewers with market shares in excess of 5%. In the late 1990s, there had been consolidation of the larger companies in the domestic beer industry. The most recent example was Stroh Brewing

Company (SBC) with about 5% market share. SBC had signed agreements to sell its major brands to Miller and the remaining brands to Pabst Brewing Company. SBC would then exit the beer industry by 2000.

From 1983 through 1998, Coors was the only major U.S. brewer to increase its sales volume each year although industry sales had grown only about 1% per year in the 1990s. Coors had outpaced the industry volume growth rate by 1% or 2% points each year. Coors had accomplished this growth by building its key premium brands in key markets and strengthening its distributor network, recently with improved supply-chain management.

Coors Vision Statement and Business Strategies. Coors vision statement was as follows:

Our company has a proud history of visionary leadership, quality products, and dedicated people, which has enabled us to succeed in a highly competitive and regulated industry. We must continue to build on this foundation and become even more effective by aligning and uniting the human, financial, and physical aspects of our company to bring great tasting beer, great brands, and superior service to our distributors, retailers, and consumers and to be a valued neighbor in our communities. Our continued success will require teamwork and an even stronger dedication by every person in our organization to a common purpose, our Vision. Achieving our Vision requires that we begin this journey immediately and with urgency for it will require significant change for us to thrive and win in our industry.

Using this vision statement, top management had decided to focus on four fundamentals: improving quality, improving service, boosting profitability, and developing employee skills. In the 1997 Coors annual report, both the CEO and the President discussed the following general business strategies or "six planks" to drive these fundamentals in the future:

1. *baseline growth*: we will profitably grow key brands and key markets;
2. *incremental growth*: we will selectively invest to grow high potential markets, channels, demographics, and brands;
3. *product quality*: we will continuously elevate consumer-perceived quality by improving taste, freshness, package integrity, and package appearance at point of purchase;
4. *distributor service*: we will significantly enhance distributor service as measured by improved freshness, less damage, increased on-time arrivals, and accurate order fill at a lower cost to Coors;
5. *productivity gains*: we will continuously lower total company costs per barrel so Coors can balance improved profitability, investments to grow

volume, market share and revenues, and funding for the resources needed to drive long-term productivity and success; and
6. *people*: we will continuously improve our business performance through engaging and developing our people.

Descriptions of Performance Measures

1. **Market share**: increase in Coors market share of the domestic beer market. Reported monthly.
2. **Barrels produced per labor hour**: total barrels packaged per labor hour worked. Reported daily.
3. **Baseline growth**: percentage of sales growth for key brands and key markets. Reported annually.
4. **Beer waste and package scrap**: waste and scrap as a percent of total production. Reported weekly.
5. **Community involvement**: number of volunteer hours per employee.
6. **Customer complaints**: total customer complaints related to taste, freshness, package integrity, appearance, and foreign objects per 100,000 barrels sold. Reported weekly.
7. **Incremental growth**: percentage of sales growth for high potential brands and markets. Reported annually.
8. **Load item accuracy**: percent beer line items shipped exact as compared to the commitment to the distributor. Reported daily.
9. **Load schedule performance**: truck or rail car loaded on time (within two hours of scheduled lead time). Reported daily.
10. **Manufacturing cost per barrel**: total plant cost (brewing materials, production labor, support labor, operating supplies, manufacturing overhead, maintenance materials, and packing materials) on a per barrel basis. Reported monthly.
11. **Net profit per barrel**: net income, excluding all special charges and special credits, on a per barrel basis. Reported monthly.
12. **New products**: number of new brands introduced each year. Reported annually.
13. **Plant productivity**: actual production hours divided by total production hours including run time, unplanned downtime and changeovers. Reported daily.
14. **Production stability**: total quantity of correct product from the beer lines as scheduled within a four-hour window as a percent of total production. Reported daily.

15. **Quality index**: weighted roll-up of component quality measures concerning plant audits, microbiology, and chemistry on a scale of 1–100. Reported quarterly.
16. **Return on invested capital**: after-tax income before interest expense and any special charges or credits divided by the sum of average total debt and shareholders' equity. Reported annually.
17. **Revenues per barrel**: total net revenues after reducing gross revenues by the excise taxes imposed by federal laws. Reported monthly.
18. **Safety (lost work case incident rate)**: total recordable cases that resulted in lost work as a ratio to total labor hours worked. Reported quarterly.
19. **Safety (total case incident rate)**: total OSHA recordable case incidents as a ratio to total number of labor hours worked. Reported quarterly.
20. **Selling, general and administrative (SG&A) cost per barrel**: all SG&A costs (includes selling, advertising, outbound transportation, distribution, and all general and administrative costs; excludes interest expenses, special charges, and income taxes) on a per barrel basis. Reported monthly.
21. **Skills**: inventory of cross-functional employee skills. Reported annually.
22. **Throughput**: total barrels packaged each period. Reported monthly.
23. **Training**: number of training hours per employee. Reported annually.
24. **Warehouse moves**: actual shipments as a percent of planned shipments. Reported Weekly.

Adapted from Coors Case: Balanced Scorecard (Grove et al., 2000).

APPENDIX B. PERFORMANCE MEASURES FOR EXPERIMENT 1, VERSION 1

No.	Performance Measure	Year 2000		Circle the number indicating importance of the measure in evaluating Coors.						Circle the number indicating your view of Coors' performance on this measure.					
		Actual	Target	1 = Not important; 6 = Critical						1 = Very poor; 6 = Very good					
1	Annual market share increase	0.04%	0.05%	1	2	3	4	5	6	1	2	3	4	5	6
2	Barrels produced per labor hour	5.0	6.0	1	2	3	4	5	6	1	2	3	4	5	6
3	Baseline growth (for key brands and markets)	2.0%	2.0%	1	2	3	4	5	6	1	2	3	4	5	6
4	Beer waste & package scrap (as a % of total production)	0.5%	0%	1	2	3	4	5	6	1	2	3	4	5	6
5	Community involvement (volunteer hours per employee annually)	25	30	1	2	3	4	5	6	1	2	3	4	5	6
6	Customer complaints (per 100,000 barrels sold)	0.06%	0.05%	1	2	3	4	5	6	1	2	3	4	5	6
7	Incremental growth (for high potential brands and markets)	2.5%	3.0%	1	2	3	4	5	6	1	2	3	4	5	6
8	Load item accuracy	93%	100%	1	2	3	4	5	6	1	2	3	4	5	6
9	Load schedule performance	55%	100%	1	2	3	4	5	6	1	2	3	4	5	6
10	Manufacturing cost per barrel	$55	$53	1	2	3	4	5	6	1	2	3	4	5	6
11	Net profit per barrel	$3	$6	1	2	3	4	5	6	1	2	3	4	5	6
12	New products (new brands introduced each year)	5	6	1	2	3	4	5	6	1	2	3	4	5	6

13	Plant productivity	76%	80%	1	2	3	4	5	6	1	2	3	4	5	6
14	Production stability	50%	100%	1	2	3	4	5	6	1	2	3	4	5	6
15	Quality index (out of 100)	85	90	1	2	3	4	5	6	1	2	3	4	5	6
16	Return on invested capital	7.1%	12.0%	1	2	3	4	5	6	1	2	3	4	5	6
17	Revenue per barrel	$100	$100	1	2	3	4	5	6	1	2	3	4	5	6
18	Safety (lost work incident rate)	0.0001	0	1	2	3	4	5	6	1	2	3	4	5	6
19	Safety (total case incident rate)	0.0002	0	1	2	3	4	5	6	1	2	3	4	5	6
20	SG&A cost per barrel	$28	$27	1	2	3	4	5	6	1	2	3	4	5	6
21	Skills (inventory of cross-functional employee skills)	6	7	1	2	3	4	5	6	1	2	3	4	5	6
22	Throughput per month (millions of barrels)	1.72	1.75	1	2	3	4	5	6	1	2	3	4	5	6
23	Training (hours per employee annually)	42	40	1	2	3	4	5	6	1	2	3	4	5	6
24	Warehouse moves (actual shipments as a % of plan)	95%	100%	1	2	3	4	5	6	1	2	3	4	5	6
	Please evaluate the overall performance of Coors Brewing Company in the year 2000. = = = >									1	2	3	4	5	6

Source: Adapted from *Coors Case: Balanced Scorecard* (Grove et al., 2000).

APPENDIX C. SELECTED SURVEY MATERIALS FOR EXPERIMENT 2

Background Information and Instructions

You are the new Vice-President of Operations of XYZ, Inc. Your first job as VP of Operations is to evaluate the 2003 performance of one of the manufacturing divisions, the Eastern Division. The evaluation of the Eastern Division will be used to evaluate the Division's General Manager, for future allocation of new capital, and for other miscellaneous purposes. The targets for 2003 were mutually agreed upon in October 2002 by the Division General Manager and your predecessor as VP of Operations. The performance measures and the 2003 targets reflected the strategic objectives of the Eastern Division and the general market conditions.

Below are the Eastern Division's performance measures with actual and targeted results for 2003. Please review the performance measures and answer the questions that follow.

XYZ, Inc. – Eastern Division Performance Measures, Version 1

Performance Measure	Year 2003	
	Actual	Target
Average lead time in days (order to delivery)	7.2	7
Customer satisfaction index (out of 100)	91	90
Defects per 1,000 units	49.5	50
Employee satisfaction index (out of 100)	91	90
Employee turnover	1.9%	2%
Free cash flow ($000)	$3,989	$5,000
Inventory turns	20	20
Labor efficiency	89.5%	90%
Manufacturing cycle time in days	4.1	4
Market share	20.2%	20%
Number of warranty claims per 1,000 shipments	39.7	40
On-time delivery	95.5%	95%
Operating income ($000)	$7,993	$10,000
Percentage involvement in voluntary quality circles	81.6%	80%
Purchase price variance – favorable (unfavorable)	1.02%	1.0%
Return on investment	9.6%	12%
Revenue growth rate	8.0%	10%
Revenue per unit	$1,000	$1,000
Safety (OSHA reportable lost time accidents)	2	2
Training hours per employee	39	40

Instructions: Please answer the following questions based on your review of the Eastern Division's performance measures for 2003.

On a scale of 1 – 10 where **1 is very poor and 10 is very good**, please circle the number that indicates your opinion of Eastern Division's 2003 performance on the following items:

Performance Measure	Performance
1. Average lead time in days	1 2 3 4 5 6 7 8 9 10
2. Customer satisfaction index	1 2 3 4 5 6 7 8 9 10
3. Defects per 1,000 units	1 2 3 4 5 6 7 8 9 10
4. Employee satisfaction index	1 2 3 4 5 6 7 8 9 10
5. Employee turnover	1 2 3 4 5 6 7 8 9 10
6. Free cash flow	1 2 3 4 5 6 7 8 9 10
7. Inventory turns	1 2 3 4 5 6 7 8 9 10
8. Labor efficiency	1 2 3 4 5 6 7 8 9 10
9. Manufacturing cycle time in days	1 2 3 4 5 6 7 8 9 10
10. Market share	1 2 3 4 5 6 7 8 9 10
11. Number of warranty claims per 1,000 shipments	1 2 3 4 5 6 7 8 9 10
12. On-time delivery	1 2 3 4 5 6 7 8 9 10
13. Operating income	1 2 3 4 5 6 7 8 9 10
14. Percentage involvement in voluntary quality circles	1 2 3 4 5 6 7 8 9 10
15. Purchase price variance	1 2 3 4 5 6 7 8 9 10
16. Return on investment	1 2 3 4 5 6 7 8 9 10
17. Revenue growth rate	1 2 3 4 5 6 7 8 9 10
18. Revenue per unit	1 2 3 4 5 6 7 8 9 10
19. Safety	1 2 3 4 5 6 7 8 9 10
20. Training hours per employee	1 2 3 4 5 6 7 8 9 10
21. Overall, how did the Eastern Division perform?	1 2 3 4 5 6 7 8 9 10

This page intentionally left blank

FINANCIAL AND NON-FINANCIAL PERFORMANCE: THE INFLUENCE OF QUALITY OF INFORMATION SYSTEM INFORMATION, CORPORATE ENVIRONMENTAL INTEGRATION, PRODUCT INNOVATION, AND PRODUCT QUALITY

Alan S. Dunk

ABSTRACT

Issues relating to the financial and non-financial performance of firms are attracting considerable research attention. Four specific factors are focused on this paper, namely quality of information system (IS) information, corporate environmental integration, product innovation, and product quality to investigate the extent to which these variables influence financial and non-financial performance. All four independent variables were found to enhance the performance assessed in non-financial terms. In contrast, the results show that product innovation alone influences

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 91–114

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14004-0

financial performance. The findings of this study suggest that the efficacy of these factors may be more effectively assessed by evaluating their impact on performance measured in non-financial terms, thereby suggesting that the inclusion of non-financial measures in performance evaluation models should enhance control system functioning.

INTRODUCTION

Research increasingly points to the need to consider performance at the organizational level from financial and non-financial perspectives, as current interest in key result indicators suggests that it is not necessarily the province of accounting numbers.[1] The perceived inadequacies of accounting-based performance measures have triggered a number of responses ranging from improved financial metrics such as economic value measures to balanced scorecards comprising financial and non-financial indicators (Ittner & Larcker, 1998a; Kaplan & Norton, 2001a, b). Non-financial measures have received considerable attention in the literature, and are widely expected to supplement those of a financial nature in assessing performance (e.g. Vaivio, 1999; Wouters, Kokke, Theeuwes, & van Donselaar, 1999; Bisbe & Otley, 2004). A primary reason for their apparent utility is that non-financial performance measures focus on activities not effectively captured by financial results (Hemmer, 1996).

Ittner and Larcker (1998b) noted that performance measurement systems play a crucial role in the development of strategic plans and in the evaluation of organizational objectives. The literature has emphasized the importance of designing performance measurement systems that capture a range of strategically important criteria in financial and non-financial terms (Lillis, 2002). However, Hoffecker and Goldenberg (1994) warned that performance measurement systems predicated on financial measures shed little light on external constituencies. As the limitations of managing solely with financial measures have been of concern for some time, a frequently held view is that focusing on financial performance indicators may not be sufficient (e.g. Collison, Grinyer, & Russell, 1996; Elnathan, Linn, & Young, 1996; Hoque & James, 2000; Kaplan & Norton, 2001a; Malina & Selto, 2001).

Kaplan and Norton (1992) proposed that financial performance measures should be complemented by those of an operational nature to obtain a better understanding of a firm's performance. Kaplan and Norton (1992)

concluded that any linkage between improvements in operational measures and financial performance is difficult to make. Making a similar point, Bartolomeo et al. (2000) argued that a firm's performance in operational areas is better measured by non-financial indicators. Ittner and Larcker (2003) indicated that firms are measuring non-financial factors such as customer loyalty and employee satisfaction with the expectation that they may have an effect on future profitability. The concerns raised by Kaplan and Norton and others suggest that the relation between non-financial measures capturing operational activities and financial performance is far from clear. Consequently, this raises the prospect that non-financial operational measures may influence performance expressed in non-financial (i.e. non-financial performance) rather than financial (i.e. financial performance) terms. Non-financial performance as defined in this study comprises elements of a non-financial nature. However, Hemmer (1996) pointed out that there is little evidence of the role of non-financial measures in the assessment of performance generally. Nevertheless, Ittner, Larcker, and Randall (2003) argued that firms may achieve enhanced performance through a greater reliance on a broad set of financial and non-financial measures.

Ittner and Larcker (1998b) stressed that the choice of performance indicators is one of the most critical challenges facing firms. Further underscoring the point, Lillis (2002) noted that although much of the prescriptive performance measurement literature is highly persuasive, implementation issues are coming to light. She indicated that many of these issues are raised in the context of the balanced scorecard because it is often considered to be the most prescriptive contemporary framework.

This study attempts to provide a theoretical articulation and empirical evidence to demonstrate the role of four specific non-financial measures that are receiving considerable literature attention in performance evaluation. Such evidence is crucial to the design of management control systems that can contribute effectively to performance assessment. Measures of quality of information system (IS) information, corporate environmental integration, product innovation, and product quality are addressed in this paper to examine the role of non-financial measures in terms of their relation to financial and non-financial performance. These measures are evaluated in this paper to investigate the role of non-financial measures in performance assessment and the rationale for their inclusion is as follows.

As considerable literature now points to the need for information quality in organizations (e.g. Nicolaou, Masoner, & Welker, 1995; Naveh & Halevy, 2000), quality of IS information is focused on in this study. Quality of IS information refers to the reliability, relevance, accuracy, precision, and

completeness of information produced by ISs (King & Grover, 1991; DeLone & McLean, 1992; Nicolaou et al., 1995).[2] Kaplan and Norton (1992) argued that ISs play a crucial role in the disaggregation of performance indicators relating to internal business process operations. Corporate environmental integration's contribution has also emerged in the literature as a critical factor to address. It refers to the extent to which firms integrate environmental issues into their management control systems and involves the identification, measurement, reporting, monitoring, and managing environmental effects through a firm's management control systems (Epstein, 1996; Klassen & Whybark, 1999). Environmental regulation has become a major organizational concern as studies report that firms typically spend between 1 and 2 percent of their revenues in response to environmental matters (Rugman & Verbeke, 1998; Moneva & Llena, 2000). Compliance with such regulation can therefore significantly affect the cost of products, and firms are under considerable pressure to better manage those costs within a framework of greater environmental accountability (e.g. Bailey, 1999; Joshi, Krishnan, & Lave, 2001; Al-Tuwaijri, Christensen, & Hughes II, 2004). Epstein (1996) proposed that corporate environmental integration is one means of doing so.

Kaplan and Norton (1992) also argued that a company's ability to innovate is linked directly to its value, and hence product innovation is an important factor in this study. Product innovation refers to the adoption of an internally generated or purchased product that is new to the adopting organization (e.g. Damanpour, 1991). Hitt and Hoskisson (1997) reported following their literature review that product innovation is often considered to be a significant contributor to a firm's competitive advantage. The remaining factor focused on in this study is product quality. Although work in accounting has often considered product quality from the cost of quality perspective (e.g. Foster & Sjoblom, 1996; Anderson & Sedatole, 1998; Nagar & Rajan, 2001), it has more generally been defined in terms of the extent to which consumers' needs and desires are reflected in products being marketed (Reeves & Bednar, 1994; Lynch, 1999). Lillis (2002) noted recently that rapid global changes in manufacturing competition have emphasized the importance of product quality to firms.

The purpose of this paper is to investigate the extent to which the quality of IS information, corporate environmental integration, product innovation, and product quality influence the financial and/or non-financial performance of firms. The results of this study, consistent with the expectations of the literature, suggest that each of these independent variables positively influence non-financial performance. The findings also show that product

innovation has a positive impact on financial performance, whereas the other three factors do not. These results indicate that the performance of firms should be assessed in both financial and non-financial terms as non-financial factors increase in prominence.

The remainder of the paper is structured as follows: The next section reviews the literature and proposes two hypotheses. The following section discusses the method used in data collection together with a psychometric analysis of the measures employed in hypothesis testing. The subsequent section presents the results. The final section addresses the conclusions drawn from the study as well as its potential limitations.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Evidence suggests that the management control systems of many firms are predicated on financial performance indicators (e.g. Kaplan & Norton, 1996a; Ittner & Larcker, 1997). For example, Coates, Davis, and Stacey (1995) reported that profitability, share price growth, return on equity, return on investment, as well as profit and earnings per share are frequent performance indicators. Similarly, Clinton and Hunton (2001) measured performance based on percentage change in net income, percentage change in stock price, percentage change in return on investment for the most recent reporting year, and a rating of overall performance as compared to peer organizations.

Kaplan and Norton (1992) argued that the link between non-financial measures capturing and reporting on operational activities and financial performance is far from clear. They drew attention to an NYSE company that was considered to have made significant improvements in quality, productivity, and customer service, which did not appear to be reflected in its financial performance. Subsequently, Kaplan and Norton (1996b) commented on the financial difficulties winners of the Baldrige quality award have experienced. Richardson, Vandenberg, Blum, and Roman (2002) recently argued that although improvements in financial performance is critical to many organizations, firms may also regard paying attention to non-financial performance factors as important. Apart from such anecdotes, Banker, Potter, and Srinivasan (2000) reported that few studies have addressed the relation between non-financial measures and financial performance. Furthermore, although an argument for the use of non-financial

measures is mounted within the balanced scorecard framework, evidence of their role and utility within that framework is limited (Otley, 1999; Vaivio, 1999; Norreklit, 2000). Hence, there is little substantive work to evidence a linkage between measures and financial or non-financial performance.

Nevertheless, Kaplan and Norton (1992) argued that non-financial measures gauge the effect of managerial actions that will not be reflected in current financial performance. Following a review of the literature, Banker et al. (2000) similarly noted that a primary reason for the utility of non-financial measures is that they may be better indicators of future financial performance than financial measures. However, in a study focusing on one non-financial measure, they reported that there was no evidence to suggest there is a systematic association between customer satisfaction and financial performance at the firm level when lead-lags are taken into consideration. Moreover, Banker et al. (2000) argued that there is little evidence of a theory to support the nature of any such lead-lag relation. Consequently, this paper focuses on examining the relation between four specific factors and the financial and non-financial performance of firms without attempting to estimate lead or lag periods, given the absence of theory to articulate them. Arguments follow to develop the relation between quality of IS information, corporate environmental integration, product innovation and product quality, and the two criterion variables.

Quality of IS Information

Although Robbins and Stylianou (1999) argued that companies depend on their information systems to facilitate the provision of accurate, reliable and up-to-date information, firms report that these systems no longer provide enduring competitive advantage, and that organizations need to shift their attention from systems to the nature of the information being generated (Yuthas & Eining, 1995; Lee, Gosain, & Im, 1999; Levy, Powell, & Galliers, 1999). The literature has similarly made it clear that firms rely on their IS to facilitate the provision of high-quality information necessary for organizational functioning (Mirani & Lederer, 1998; Robbins & Stylianou, 1999; Naveh & Halevy, 2000). Reports suggest that the value relevance of that information depends on its contribution to the control of organizational activities and to the role it plays in supporting decision making relating to the products and services the firms deliver (e.g. Teng, Cheon, & Grover, 1995). For example, firms stress that their performance depends in part on high quality information as it provides them with a basis for

strategic advantage (Mukherjee, Lapre, & Wassenhove, 1998; Typanski, 1999; Lainhart IV, 2000).

Propositions such as these have led to greater attention being placed on the quality of IS information as a means of promoting performance through its contribution to improvements in an array of decision processes. Consequently, the literature suggests that the quality of IS information is more likely to promote non-financial than financial performance (e.g. Teng et al., 1995; Naveh & Halevy, 2000).

Corporate Environmental Integration

Corporate environmental integration facilitates the setting of environmental goals and the means of their implementation and attainment (Epstein, 1996; Brady, Henson, & Fava, 1999). Hunt and Auster (1990) argued that organizations need comprehensive environmental management systems, and corporate environmental integration is one such system. A possible benefit of having a system of corporate environmental integration is that it may facilitate a firm's capacity to manage its environmental accountability in response to societal constraints (Al-Tuwaijiri et al., 2004). Proposals also indicate that the manner in which organizations address environmental issues has the potential to directly affect the marketability of their products and their competitive position (e.g. Post & Altman, 1992; Billing & Scott, 1995). For example, Porter and van der Linde (1995) emphasized that the integration of environmental matters into management systems facilitates environmentally sensitive product and process improvements. Epstein (1996) found that as companies integrate environmental factors into product costing, capital investment, and performance evaluation systems, they typically report that performance improves. Such performance enhancements are reportedly evidenced particularly by improvements in a firm's corporate and product image, as well as in its market share for green products (White, Becker, & Savage, 1993; Klassen & Whybark, 1999). Hence, the identification and management of environmental impacts through corporate environmental integration should result in an increased non-financial performance (Zhang, Kuo, Lu, & Huang, 1997; Judge & Douglas, 1998; Brady et al., 1999; Moneva & Llena, 2000).

Product Innovation

A clear consensus has emerged whereby product innovation is characterized as being critically important to a firm's performance in an increasingly

competitive marketplace (e.g. Bromwich, 1990; Balkin, Markman, & Gomez-Mejia, 2000). Competitiveness in world markets frequently depends in part on the ability of firms to develop and market innovative products (Shields & Young, 1994; Tijssen & van Wijk, 1999; Meade & Presley, 2002). Product innovation is used by firms to enter new industries, to market entirely new products, and to gain advantage over competitors (Ali, 1994; Greve & Taylor, 2000). Consequently, the literature regards product innovation as an important contributor to a firm's operations (e.g. Ettlie & Reza, 1992; Hitt & Hoskisson, 1997; Li & Atuahene-Gima, 2001).

The literature suggests that product innovation is likely to be linked to a firm's financial performance in addition to the expectation that it influences non-financial performance (e.g. Capon, Farley, Lehmann, & Hulbert, 1992; Shields & Young, 1994; Balkin et al., 2000). Baer and Frese (2003) recently argued that innovation is frequently regarded by firms as crucial to increasing profits and market share. One rationale for product innovation influencing financial performance is that it has become important for value creation in many firms (Hitt, Hoskisson, Johnson, & Moesel, 1996). Arguments have also been raised that innovative products often allow firms to command price premiums, which are likely to promote profitability (Porter, 1990; Ali, 1994). Consequently, innovative products are expected to improve a firm's financial and non-financial performance (De Maio, Verganti, & Corso, 1994; Calantone, Vickery, & Droger, 1995; Nijssen, Arboun, & Commandeur, 1995).

Product Quality

Enhancing product quality has been considered a strategic organizational priority for some time. Flynn, Schroeder, and Sakakibara (1994) argued that quality is a critical component in the design and manufacture of products, which are superior to those of competitors, and therefore it has a role to play in influencing performance. Arguments also suggest that quality provides a basis for strategic advantage and thus improvements in product quality should lead to performance improvements (e.g. Daniel & Reitsperger, 1991; Belohlav, 1993; Terziovski, Sohal, & Moss, 1999).

However, discussion in the literature supports the proposal that product quality does not directly influence financial performance. For example, Hitt and Hoskisson (1997) argued that as customers increasingly expect products to be of high quality, product quality is unlikely to have a direct impact on a firm's financial performance. Furthermore, Nagar and Rajan

(2001) reported there is a widespread concern that traditional financial measures of quality do not capture product quality. Critics of traditional quality cost systems propose supplementing financial measures with non-financial quality measures, arguing that non-financial measures provide a better indication of quality-related outcomes (Ittner & Larcker, 1997; Nagar & Rajan, 2001). However, Norreklit (2000) warned that the relation between non-financial measures of product quality and financial performance has not been empirically demonstrated. More recently, Nagar and Rajan (2001) similarly proposed that non-financial quality measures better capture the effects of product quality. Consequently, the literature suggests that product quality is more likely to positively influence non-financial than financial performance.

Hypotheses

The review of the literature suggests that quality of IS information, corporate environmental integration, product innovation, and product quality enhance a firm's performance measured in non-financial terms. In contrast, the literature indicates that only product innovation is likely to positively affect financial performance. These propositions are expressed in the following hypotheses stated in alternate form:

H1. Quality of IS information, corporate environmental integration, product innovation, and product quality positively affect non-financial performance.

H2. Quality of IS information, corporate environmental integration, and product quality do not affect financial performance, whereas product innovation does.

METHOD

A random sample of 119 functional area managers was drawn from manufacturing organizations across Australia listed in *Kompass Australia* employing more than 100 people. Each manager was contacted by telephone and requested to take part in the study. On agreeing to do so, each manager was mailed an anonymous questionnaire together with a cover letter and a stamped addressed envelope for its return. A telephone follow-up was conducted 2 weeks later to enhance the response rate. The follow-up also

provided considerable assurance that the managers had completed the questionnaire themselves.

A total of 77 managers responded, representing a response rate of 65 percent. The sample comprised 26 marketing and 42 production managers, together with 9 other managers from a range of areas of responsibility. Their average age was 43 and the mean years of experience in the areas they managed was 12. They had held their present positions on an average of 4 years and the mean number of employees in their areas of responsibility was 85. Firms sampled were involved in the manufacture of pharmaceuticals, white goods, beverages, foodstuffs, chemicals, farm equipment, as well as information technology equipment, automotive, and building products. Incomplete responses were received in two instances to the quality of IS information scale, one each to the product innovation, financial and non-financial performance instruments, and three to the product quality scale. Consequently, statistical analyses are based on a sample size of 70.[3]

Variable Measurement

Quality of IS Information

Quality of IS information was measured using the Teng et al. (1995) five-item, seven-point Likert-scaled fully anchored instrument. The meaning of information's attributes have long been in place, for example, reliability refers to information that can be depended on, relevance means that it is important to the decision maker, and accuracy indicates that it is error-free (e.g. Wang & Strong, 1996). Consequently, respondents were asked to rate their IS in terms of its reliability, relevancy, accuracy, precision, and completeness on a scale anchored by (1) very low and (7) very high. Descriptive statistics for the measure are presented in Table 1.

Table 2 shows that all items load on a single factor having an eigenvalue of 3.603, explaining 72.1 percent of the variance in the underlying variable. The Cronbach alpha of 0.927 indicates that the internal consistency of the instrument is high and is consistent with that of Teng et al. (1995), who reported a Cronbach alpha for the scale of 0.886, with all items loading on a single factor.

Corporate Environmental Integration

Corporate environmental integration was measured by an eight-item, seven-point Likert-scaled instrument based on Epstein's (1996) corporate

Table 1. Descriptive Statistics of the Variables in the Study.

Variable	*n*	Mean	Std. Dev.	Theoretical		Actual	
				Min	Max	Min	Max
Quality of IS information	75	22.440	5.891	5	35	7	35
Corporate environmental integration	77	34.210	9.960	8	56	11	56
Product innovation	76	24.618	7.343	6	42	8	38
Product quality	74	19.946	4.545	4	28	10	28
Financial performance	76	14.395	3.997	3	21	6	21
Non-financial performance	76	22.395	4.836	5	35	13	32

Table 2. Factor Analysis of Quality of IS Information.

Item	Factor Loading	Eigenvalue	Percent of Variance
Accuracy	0.874		
Precision	0.801		
Reliability	0.816		
Completeness	0.911		
Relevancy	0.838	3.603	72.1

environmental integration scorecard. Respondents were asked to indicate the extent to which environmental issues have been incorporated into each of the eight items including corporate strategy, product design, and capital budgeting systems, based on a scale anchored by (1) to no extent and (7) to a great extent. A factor analysis, as shown in Table 3, indicates that all items load on a single factor with an eigenvalue of 4.922, explaining 61.5 percent of the variance in the underlying variable. The Cronbach alpha of 0.925 for the scale suggests that its internal consistency is high. Descriptive statistics for the measure are shown in Table 1.

Product Innovation

Miller and Friesen's (1982) six-item, seven-point Likert-scaled instrument was used to measure product innovation. Respondents indicated on a scale anchored by (1) strongly disagree and (7) strongly agree, the extent to which they agreed with each of the items with respect to their company. Examples

Table 3. Factor Analysis of Corporate Environmental Integration.

Item	Factor Loading	Eigenvalue	Percent of Variance
Corporate strategy	0.620		
Product design systems	0.704		
ISs for internal reporting	0.805		
Internal auditing systems	0.770		
External reporting	0.788		
Costing systems	0.881		
Capital budgeting systems	0.866		
Performance evaluation systems	0.809	4.922	61.5

Table 4. Factor Analysis of Product Innovation.

Item	Factor Loading	Eigenvalue	Percent of Variance
Emphasis on the marketing of tried and true products (R)	0.351		
Emphasis on R&D, technological leadership, and innovations	0.537		
We have marketed very few lines of product in the past 5 years (R)	0.835		
We have marketed a great number of new product lines in the past 5 years	0.807		
Changes in our product lines have been mostly of a minor nature (R)	0.762		
Changes in our product lines have usually been dramatic	0.740	2.889	48.1

Note: (R) indicates reverse-scaled items.

of such items are first, there is a strong emphasis on the marketing of tried and true products and second, changes in product lines have been mostly of a minor nature. The results of a factor analysis, as shown in Table 4, indicate that all items load on a single factor with an eigenvalue of 2.889, explaining 48.1 percent of the variance in the underlying variable. The Cronbach alpha of 0.836 found in this study, and consistent with that of Miller and Friesen's (1982) alpha of 0.770, suggests that the internal consistency of the scale is relatively high. Descriptive statistics for the instrument are presented in Table 1.

Product Quality

Product quality was measured by the Flynn et al. (1994) four-item, seven-point Likert-scaled instrument. Respondents were asked to indicate on a scale anchored by (1) strongly disagree and (7) strongly agree, the extent to which they agreed with each of the items in relation to their company. Examples of scale items are, whether new product designs are reviewed before the product is produced and sold and whether customer requirements are analyzed in the new product design process. The results of a factor analysis, shown in Table 5, indicate that all items load on a single factor with an eigenvalue of 1.704, explaining 42.6 percent of the variance in the underlying variable. The Cronbach alpha of 0.727, consistent with that of Flynn et al. (1994) of 0.723, suggests that the internal consistency of the scale is satisfactory. Table 1 presents descriptive statistics for the measure.

Financial and Non-Financial Performance

To measure financial and non-financial performance, the Govindarajan and Fisher (1990) instrument was used, also previously employed by Govindarajan (1988). The instrument was selected for its diverse array of financial

Table 5. Factor Analysis of Product Quality.

Item	Factor Loading	Eigenvalue	Percent of Variance
New product designs are thoroughly reviewed before the product is produced and sold	0.797		
Customer requirements are thoroughly analyzed in the new product design process	0.771		
Reducing the cost of new products is a more important priority than new product quality (R)	0.500		
On-time delivery concerns are more important than quality in the new product development process (R)	0.474	1.704	42.6

Note: (R) indicates reverse-scaled items.

and non-financial performance indicators. Bisbe and Otley (2004, p. 719) also used this measure, noting that it is a "well-established multi-dimensional instrument". Managers were asked to rate the performance of their firms on a ten-item, seven-point Likert scale anchored by (1) well below average and (7) well above average. Following a varimax rotation, the items were found to load on three factors as shown in Table 6. The table indicates that factor 1 comprises the three items relating to profit, cash flow from operations, and return on investment, with an eigenvalue of 4.067, explaining 40.7 percent of the variance in the underlying variable. These three items are specifically referred to by Kaplan and Norton (1993) as measures of financial performance. Factor 2 incorporates the non-financial performance items of new product development, new market development, cost management, personnel development, and political/public affairs.[4] The eigenvalue for this factor is 1.274, explaining 12.7 percent of the variance in the underlying variables. Factor 3 is made up of the two items of sales volume and market share, and as the eigenvalue for this factor is less than one (0.801), it formed no part of the hypothesis testing procedures (Rummel, 1970). The Cronbach alphas for factors 1 and 2 are 0.905 and 0.775, respectively. Descriptive statistics for the measures are shown in Table 1.[5,6,7]

Table 6. Factor Analysis of Organizational Performance: Financial and Non-Financial Dimensions.

Item	Factor Loadings		
	Factor 1	Factor 2	Factor 3
Sales volume	0.235	0.228	0.913
Market share	0.391	0.081	0.650
Profit	0.882	0.209	0.193
Cash flow from operations	0.742	0.362	0.311
Return on investment	0.814	0.119	0.205
New product development	0.117	0.570	0.275
New market development	0.020	0.731	0.267
Cost management	0.217	0.504	0.135
Personnel development	0.130	0.705	0.087
Political/public affairs	0.204	0.600	0.043
Eigenvalue	4.067	1.274	0.801
Percent of variance	40.7	12.7	8.0

Table 7. Correlation Matrix of the Variables in the Study.

Variable	Financial Performance	Non-Financial Performance	Quality of IS Information	Corporate Environmental Integration	Product Innovation
Non-financial performance	0.424***				
Quality of IS information	0.315**	0.479***			
Corporate environmental integration	0.319***	0.510***	0.288**		
Product innovation	0.300**	0.451***	0.136	0.182**	
Product quality	0.246**	0.554***	0.243**	0.277**	0.219

**$p<0.05$.
***$p<0.01$.

Correlation Matrix

Table 7 provides a correlation matrix of the independent and dependent variables in the study. The magnitude of the intercorrelations between the predictors indicates that multicollinearity would not affect the results of the research (Lewis-Beck, 1990). Moreover, the correlation between financial and non-financial performance does not suggest that the criterion variables are not independent, as the factor analysis rotation was orthogonal (Kerlinger, 1986). Moreover, Williams, Macintosh, and Moore (1990) noted that since the items comprising any one factor do not load highly or exclusively on that factor, then correlations between factors can be non-zero, despite the use of varimax rotation. Therefore, items that do not load uniquely on a factor contribute to factor intercorrelation.

RESULTS

The following equation was used in testing the hypotheses:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + e \quad (1)$$

where Y is the non-financial (H1), and financial performance (H2), X_1 the quality of IS information; X_2 the corporate environmental integration; X_3 the product innovation; and X_4 the product quality.

Table 8. Results of Hypothesis Test: Non-Financial Performance.

Variable	Coefficient	Value	Std. Error	t	p
Constant	b_0	1.595	2.183	0.73	0.468
Quality of IS information	b_1	0.202	0.065	3.09	0.003
Corporate environmental integration	b_2	0.132	0.040	3.32	0.001
Product innovation	b_3	0.164	0.053	3.10	0.003
Product quality	b_4	0.401	0.086	4.68	0.001

Notes: Adjusted $R^2 = 0.576$, $n = 70$, $F_{4,65} = 24.44$, $p < 0.001$.

Table 9. Results of Hypothesis Test: Financial Performance..

Variable	Coefficient	Value	Std. Error	t	p
Constant	b_0	4.795	2.609	1.84	0.071
Quality of IS information	b_1	0.125	0.078	1.60	0.114
Corporate environmental integration	b_2	0.076	0.047	1.61	0.112
Product innovation	b_3	0.112	0.063	1.78	0.080
Product quality	b_4	0.086	0.103	0.84	0.407

Notes: Adjusted $R^2 = 0.147$, $n = 70$, $F_{4,65} = 3.98$, $p = 0.006$.

Two regression analyses were run, one in which the dependent variable is non-financial performance (H1), and the other in which financial performance is the criterion variable (H2). The results of these analyses are presented in Tables 8 and 9. Table 8 reveals that all four independent variables affect performance assessed from a non-financial perspective. Quality of IS information positively influences performance ($t = 3.09$, $p = 0.003$), as does corporate environmental integration ($t = 3.32$, $p = 0.001$), product innovation ($t = 3.10$, $p = 0.003$), and product quality ($t = 4.68$, $p = 0.001$). Hence, these results provide strong support for H1. The four predictors explain a total of 57.6 percent of the variance in non-financial performance.[8]

The findings relating to H2 presented in Table 9 illustrate that product innovation positively affects financial performance ($t = 1.78$, $p = 0.080$). However, the coefficients of corporate environmental integration ($t = 1.61$, $p = 0.112$), product quality ($t = 0.84$, $p = 0.407$), and quality of IS information ($t = 1.60$, $p = 0.114$) were found not to be different from zero. Hence, H2 is also supported.[9]

CONCLUSIONS

The results of this study suggest that the quality of IS information, corporate environmental integration, product innovation, and product quality positively influence non-financial performance, consistent with the expectations of H1. Given the literature interest in these four predictors, and the magnitude of the variance they explain in the level of non-financial performance, the findings of this research suggest that they are critical variables to consider in the assessment of performance. Furthermore, with respect to H2, the results of this study show that product innovation also affects financial performance. However, as H2 expected, the coefficients of quality of IS information, corporate environmental integration, and product quality failed to reach significance.

Consequently, these findings are a matter of some interest in terms of the impact of the latter three factors on the performance of firms from a financial perspective. The results of this study suggest that the specific inclusion of non-financial measures in performance evaluation models complement such control system focusing.

A number of limitations may have influenced the results of this study. Data were drawn only from firms in Australia and hence the results may only be generalizable to that population. The use of a self-rating measure of performance may be regarded as a limitation of this research, yet it is far from clear what would constitute an alternative comprehensive measure of performance. Furthermore, as the research is cross-sectionally based, no statement of causation can be made. In addition, any time series aspects of the study in terms of the impact of the predictors on the dependent variables cannot be ascertained. Further research is needed to investigate the potential impact of lead or lag periods, given the absence of theory or evidence to articulate such periods (Banker et al., 2000). As prior reliability and validity coefficients for the corporate environmental integration instrument do not exist, further psychometric analysis would be advantageous.

NOTES

1. Performance regarded in this paper is consistent with the financial/non-financial dichotomy frequently used in the literature (e.g. McNair, Lynch, & Cross, 1990; Banker et al., 2000; Bisbe & Otley, 2004).

2. The proliferation of sophisticated computerized information systems has increased the availability of information (Teng et al., 1995; Tuttle & Kershaw, 1998). Reference to IS in the information quality term makes the point clear.

3. There are no differences in the distribution of independent and dependent variables scores for marketing and production managers (quality of IS information: $t = 0.36$, n.s., corporate environmental integration: $t = 1.51$, n.s., product innovation: $t = 1.35$, n.s., product quality: $t = 1.24$, n.s., financial performance: $t = 0.32$, n.s., and non-financial performance: $t = 0.26$, n.s.). Hence, their scores may be aggregated.

4. Cost management is non-financial in this context as the item refers specifically to the extent to which it has been incorporated in an organization's management system.

5. The items comprising the four independent variables were simultaneously subjected to a varimax-rotated factor analysis, and the results indicated that the items comprising each variable load on each respective factor, based on the decision heuristic cutoff of at least 0.400 (Rummel, 1970). These results support factor independence.

6. To ensure whether any relation between the dependent and independent variables was due to any underlying item similarity, a series of factor analyses were conducted. The results indicated that none of the independent variable items loaded on a dependent variable factor was greater than 0.256. Kim and Mueller (1978) indicated that factor loadings less than 0.3 were insubstantial.

7. Williams et al. (1990) argued that self-rated performance scores overcome the difficulty of matching data objectively across organizations, and avoid the halo effects associated with superiors' ratings. Furthermore, even though managers may overrate their performance, or their areas of responsibility, Venkatraman and Ramanujan (1987) found that managers' self-ratings are less biased than researchers might expect. Abernethy and Stoelwinder (1991) also found no evidence that managers are consistently lenient when rating performance. Concerns for possible bias in performance ratings may also be evaluated from the perspective of measurement error, classifiable as non-random (systematic bias) or random (Nunnally, 1981; Carmines & Zellar, 1982). Although bias in criterion variable measurement changes the intercept of a regression of performance, it does not affect the other regression coefficients (Nunnally, 1981). In contrast, random error is absorbed in the model error term (Neter, Wasserman, & Kutner, 1985), thereby reducing the likelihood of the null hypothesis being rejected.

8. To assess whether performance in financial terms impacts performance from a non-financial perspective, financial performance was added as an additional independent variable in a regression in which non-financial performance was the dependent variable. The results showed that the quality of IS information ($t = 2.83$, $p = 0.006$), corporate environmental integration ($t = 3.05$, $p = 0.003$), product innovation ($t = 2.80$, $p = 0.007$), and product quality ($t = 4.55$, $p < 0.001$) remained significant, whereas financial performance was not ($t = 1.04$, $p = 0.301$). The significance of the four predictors in this model is consistent with those of the original analysis reported in Table 8.

9. To examine whether the assumptions of regression analysis had been contravened in using Eq. (1) to assess the hypotheses, separate tests were conducted to assess the homogeneity of variance of the residuals, the appropriateness of a linear model, and the normality of residuals (Cohen & Cohen, 1983). The plot of the standardized residuals against each dependent variable revealed no evidence that the

assumption of homoscedasticity had been violated. Plots of the residuals against each of the predictors in each model indicated that the data points all fell within a horizontal band of uniform width about a zero mid-point, suggesting that a linear relation represented the effects of each predictor. A plot of the residuals and their normal scores indicated that there was no departure from normality. The positive correlations ($r = 0.987$, $p < 0.001$ and $r = 0.992$, $p < 0.001$ for the financial and non-financial performance hypothesis tests, respectively) between the standardized residuals and the normal scores provided support for the normality of the residuals (Ryan, Joiner, & Ryan, 1982).

ACKNOWLEDGMENTS

Thanks are due to Carlin Dowling, Jack English, Marc Epstein, Sue Hrasky, Errol Iselin, Lokman Mia, Chew Ng, as well as to the participants at seminars at Bond University, the University of Canberra, Griffith University, the University of Tasmania, and to those at a session of the 2003 annual congress of the European Accounting Association. The helpful comments of the anonymous reviewer are also greatly appreciated.

REFERENCES

Abernethy, M. A., & Stoelwinder, J. U. (1991). Budget use, task uncertainty, system goal orientation and subunit performance: A test of the 'fit' hypothesis in not-for-profit hospitals. *Accounting, Organizations and Society*, *16*(2), 105–120.

Ali, A. (1994). Pioneering versus incremental innovation: Review and research propositions. *Journal of Product Innovation Management*, *11*(1), 46–61.

Al-Tuwaijri, S. A., Christensen, T. E., & Hughes, K. E., II (2004). The relations among environmental disclosure, environmental performance, and economic performance: A simultaneous equations approach. *Accounting, Organizations and Society*, *29*, 447–471.

Anderson, S. W., & Sedatole, K. (1998). Designing quality into products: The use of accounting data in new product development. *Accounting Horizons*, *12*(3), 213–233.

Baer, M., & Frese, M. (2003). Innovation is not enough: Climates for initiative and psychological safety, process innovations, and firm performance. *Journal of Organizational Behavior*, *24*(1), 45–68.

Bailey, I. G. (1999). Competition, sustainability and packaging policy in the UK. *Journal of Environmental Planning and Management*, *42*(1), 83–102.

Balkin, D. B., Markman, G. D., & Gomez-Mejia, L. R. (2000). Is CEO pay high in high-technology firms related to innovation? *Academy of Management Journal*, *43*(6), 1118–1129.

Banker, R. D., Potter, G., & Srinivasan, D. (2000). An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review*, *75*(1), 65–92.

Bartolomeo, M., Bennett, M., Bouma, J. J., Heydkamp, P., James, P., & Wolters, T. (2000). Environmental management accounting in Europe: Current practice and future potential. *The European Accounting Review*, *9*(1), 31–52.

Belohlav, J. A. (1993). Quality, strategy, and competitiveness. *California Management Review*, *35*(3), 55–67.

Billing, R., & Scott, B. (1995). Renewable reporting. *CA Magazine*, *128*(2), 62–64.

Bisbe, J., & Otley, D. (2004). The effects of the interactive use of management control systems on product innovation. *Accounting, Organizations and Society*, *29*, 709–737.

Brady, K., Henson, P., & Fava, J. A. (1999). Sustainability, eco-efficiency, life-cycle management and business strategy. *Environmental Quarterly Management*, *8*(3), 33–41.

Bromwich, M. (1990). The case for strategic management accounting: The role of accounting information for strategy in competitive markets. *Accounting, Organizations and Society*, *15*(1/2), 27–46.

Calantone, R. J., Vickery, S. K., & Droger, C. (1995). Business performance and strategic new product development activities: An empirical investigation. *Journal of Product Innovation Management*, *12*, 214–223.

Capon, N., Farley, J. U., Lehmann, D. R., & Hulbert, J. M. (1992). Profiles of product innovators among large U.S. manufacturers. *Management Science*, *38*(2), 157–169.

Carmines, E. G., & Zellar, R. A. (1982). *Reliability and Validity Assessment*. Beverly Hills, CA: Sage.

Clinton, B. D., & Hunton, J. E. (2001). Linking participative budgeting congruence to organization performance. *Behavioral Research in Accounting*, *13*, 127–141.

Coates, J., Davis, T., & Stacey, R. (1995). Performance measurement systems, incentive reward schemes and short-termism in multinational companies: A note. *Management Accounting Research*, *6*, 125–135.

Collison, D. J., Grinyer, J. R., & Russell, A. (1996). UK managers' decisions and their perceptions of capital markets. *Journal of International Accounting and Taxation*, *5*(1), 39–52.

Damanpour, F. (1991). Organizational innovation: A meta-analysis of determinants and moderators. *Academy of Management Journal*, *34*(3), 555–590.

Daniel, S. J., & Reitsperger, W. D. (1991). Linking quality strategy with management control systems: Empirical evidence from Japanese industry. *Accounting, Organizations and Society*, *16*(7), 601–618.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95.

De Maio, A. R., Verganti, R., & Corso, M. (1994). A multi-project management framework for new product development. *European Journal of Operational Research*, *78*, 178–191.

Elnathan, D., Linn, T. W., & Young, S. M. (1996). Benchmarking and management accounting: A framework for research. *Journal of Management Accounting Research*, *8*, 37–54.

Ettlie, J. E., & Reza, E. M. (1992). Organizational integration and process innovation. *Academy of Management Journal*, *35*(4), 795–827.

Epstein, M. J. (1996). *Measuring corporate environmental performance: Best practices for costing and managing an effective environmental strategy*. Chicago, IL: Irwin.

Flynn, B. B., Schroeder, R. G., & Sakakibara, S. (1994). A framework for quality management research and an associated measurement instrument. *Journal of Operations Management*, *11*(4), 339–366.

Foster, G., & Sjoblom, L. (1996). Quality improvement drivers in the electronics industry. *Journal of Management Accounting Research*, *8*, 55–86.

Govindarajan, V. (1988). A contingency approach to strategy implementation at the business-unit level: Integrating administrative mechanisms with strategy. *Academy of Management Journal*, *31*(4), 828–853.

Govindarajan, V., & Fisher, J. (1990). Strategy, control systems, and resource sharing: Effects on business-unit performance. *Academy of Management Journal*, *33*(2), 259–285.

Greve, H. R., & Taylor, A. (2000). Innovations as catalysts for organizational change: Shifts in organizational cognition and search. *Administrative Science Quarterly*, *45*(1), 54–80.

Hemmer, T. (1996). On the design and choice of "modern" management accounting measures. *Journal of Management Accounting Research*, *8*, 87–116.

Hitt, M. A., & Hoskisson, R. E. (1997). International diversification: Effects on innovation and firm performance in product-diversified firms. *Academy of Management Journal*, *40*(4), 767–798.

Hitt, M. A., Hoskisson, R. E., Johnson, R. A., & Moesel, D. D. (1996). The market for corporate control and firm innovation. *Academy of Management Journal*, *39*(5), 1084–1119.

Hoffecker, J., & Goldenberg, C. (1994). Using the balanced scorecard to develop companywide performance measures. *Journal of Cost Management*, *8*(3), 5–17.

Hoque, Z., & James, W. (2000). Linking balanced scorecard measures to size and market factors: Impact on organizational performance. *Journal of Management Accounting Research*, *12*, 1–17.

Hunt, C. B., & Auster, E. R. (1990). Proactive environmental management: Avoiding the toxic trap. *Sloan Management Review*, *7*(Winter), 7–18.

Ittner, C. D., & Larcker, D. F. (1997). Quality strategy, strategic control systems, and organizational performance. *Accounting, Organizations and Society*, *22*(3/4), 293–314.

Ittner, C. D., & Larcker, D. F. (1998a). Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research*, *10*, 205–238.

Ittner, C. D., & Larcker, D. F. (1998b). Are financial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research*, *36*(Suppl.), 1–35.

Ittner, C. D., & Larcker, D. F. (2003). Coming up short. *Harvard Business Review*, *81*(11), 88–95.

Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance implications of strategic performance measurement in financial services firms. *Accounting, Organizations and Society*, *28*, 715–741.

Joshi, S., Krishnan, R., & Lave, L. (2001). Estimating the hidden costs of environmental regulation. *The Accounting Review*, *81*(11), 171–198.

Judge W. Q., Douglas,T. J. (1998). Performance implications of incorporating natural environmental issues into the strategic planning process: An empirical assessment. *Journal of Management Studies*, *35*(2), 241–262.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard – measures that drive performance. *Harvard Business Review*, *70*(1), 71–79.

Kaplan, R. S., & Norton, D. P. (1993). Putting the balanced scorecard to work. *Harvard Business Review*, *71*(5), 134–147.
Kaplan, R. S., & Norton, D. P. (1996a). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, *74*(1), 75–85.
Kaplan, R. S., & Norton, D. P. (1996b). *The balanced scorecard: Translating strategy into action*. Boston, MA: Harvard Business School Press.
Kaplan, R. S., & Norton, D. P. (2001a). Transforming the balanced scorecard from performance measurement to strategic management: Part I. *Accounting Horizons*, *15*(1), 87–104.
Kaplan, R. S., & Norton, D. P. (2001b). Transforming the balanced scorecard from performance measurement to strategic management: Part II. *Accounting Horizons*, *15*(2), 147–160.
Kerlinger, F. N. (1986). *Foundations of behavioral research*. Orlando, FL: Harcourt Brace Jovanovich.
Kim, J-O., & Mueller, C. W. (1978). *Factor analysis: Statistical methods and practical issues*. Beverly Hills, CA: Sage.
King, W. R., & Grover, V. (1991). The strategic use of information resources: An exploratory study. *IEEE Transactions on Engineering Management*, *38*(4), 293–305.
Klassen, R. D., & Whybark, D. C. (1999). The impact of environmental technologies on manufacturing performance. *Academy of Management Journal*, *42*(6), 599–615.
Lainhart, J. W. ,, IV. (2001). CobiTTM: A methodology for managing and controlling information and information technology risks and vulnerabilities. *Journal of Information Systems*, *14*(Suppl.), 21–25.
Lee, Z., & Gosain, S. Im,, I. (1999). Topics of interest in IS: Evolution of themes and differences between research and practice. *Information and Management*, *36*(5), 233–246.
Levy, M., Powell, P., & Galliers, R. (1999). Assessing information systems strategy development frameworks in SMEs. *Information and Management*, *36*(5), 247–261.
Lewis-Beck, M. S. (1990). *Applied regression: An introduction*. Newbury Park, CA: Sage.
Li, H., & Atuahene-Gima, K. (2001). Product innovation strategy and the performance of new technology ventures in China. *Academy of Management Journal*, *44*(6), 1123–1134.
Lillis, A. M. (2002). Managing multiple dimensions of manufacturing performance – an exploratory study. *Accounting, Organizations and Society*, *27*, 497–529.
Lynch, D. (1999). Focus on quality. *Management Accounting*, *77*(8), 30–31.
Malina, M. A., & Selto, F. H. (2001). Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research*, *13*, 47–90.
McNair, C. J., Lynch, R. L., & Cross, K. F. (1990). Do financial and nonfinancial performance measures have to agree? *Management Accounting*, *72*(5), 28–36.
Meade, L. M., & Presley, A. (2002). R&D project selection using the analytic network process. *IEEE Transactions on Engineering Management*, *49*(1), 59–66.
Miller, D., & Friesen, P. H. (1982). Innovation in conservative and entrepreneurial firms: Two models of strategic momentum. *Strategic Management Journal*, *3*, 1–25.
Mirani, R., & Lederer, A. L. (1998). An instrument for assessing the organizational benefits of IS projects. *Decision Sciences*, *29*(4), 803–838.
Moneva, J. M., & Llena, F. (2000). Environmental disclosures in the annual reports of large companies in Spain. *The European Accounting Review*, *9*(1), 7–29.
Mukherjee, A. S., Lapre, M. A., & Wassenhove, L. N. (1998). Knowledge driven quality improvement. *Management Science*, *44*(11), S35–S49.

Nagar, V., & Rajan, M. V. (2001). The revenue implications of financial and operational measures of product quality. *The Accounting Review*, *76*(4), 495–513.

Naveh, E., & Halevy, A. (2000). A hierarchical framework for a quality information system. *Total Quality Management*, *11*(1), 87–111.

Neter, J., Wasserman, W., & Kutner, M. H. (1985). *Applied linear statistical models*. Homewood, IL: Irwin.

Nicolaou, A. I., Masoner, M. M., & Welker, R. B. (1995). Intent to enhance information systems as a function of system success. *Journal of Information Systems*, *9*(2), 93–108.

Nijssen, E. J., Arboun, A. R. L., & Commandeur, H. R. (1995). Accelerating new product development: A preliminary empirical test of a hierarchy of implementation. *Journal of Product Innovation Management*, *12*, 99–109.

Norreklit, H. (2000). The balance on the balanced scorecard – a critical analysis of some of its assumptions. *Management Accounting Research*, *11*, 65–88.

Nunnally, J. C. (1981). *Psychometric theory*. New Delhi: Tata McGraw-Hill.

Otley, D. (1999). Performance management: A framework for management control systems research. *Management Accounting Research*, *10*, 363–382.

Porter, M. E. (1990). *The competitive advantage of nations*. London: Macmillan.

Porter, M. E., & van der Linde, C. (1995). Toward a new conception of the environment – competitiveness relationship. *Journal of Economic Perspectives*, *9*(4), 97–118.

Post, J. E., & Altman, B. W. (1992). Models of corporate greening: how corporate social policy and organizational learning inform leading-edge environmental management. *Research in Corporate Social Performance and Policy*, *13*, 3–29.

Reeves, C. A., & Bednar, D. A. (1994). Defining quality: Alternatives and implications. *Academy of Management Review*, *19*(3), 419–445.

Richardson, H. A., Vandenberg, R. J., Blum, T. C., & Roman, P. M. (2002). Does decentralization make a difference for the organization? An examination of the boundary conditions circumscribing decentralized decision-making and organizational financial performance. *Journal of Management*, *28*(2), 217–244.

Robbins, S. S., & Stylianou, A. C. (1999). Post-merger systems integration: The impact on IS capabilities. *Information and Management*, *36*(4), 205–212.

Rugman, A. M., & Verbeke, A. (1998). Corporate strategies and environmental regulations: An organizing framework. *Strategic Management Journal*, *19*, 363–375.

Rummel, R. J. (1970). *Applied factor analysis*. Evanston: Northwestern University Press.

Ryan, T. A., Joiner, B. L., & Ryan, B. F. (1982). *Minitab reference manual*. Pen: Ryan.

Shields, M. D., & Young, S. M. (1994). Managing innovation costs: A study of cost conscious behavior by R&D professionals. *Journal of Management Accounting Research*, *6*(Fall), 175–196.

Teng, J. T. C., Cheon, M. J., & Grover, V. (1995). Decisions to outsource information systems functions: Testing a strategy – theoretic discrepancy model. *Decision Sciences*, *26*(1), 75–103.

Terziovski, M., Sohal, A., & Moss, S. (1999). Longitudinal analysis of quality management practices in Australian organizations. *Total Quality Management*, *10*(6), 915–926.

Tijssen, R. J. W., & van Wijk, E. (1999). In search of the European paradox: An international comparison of Europe's scientific performance and knowledge flows in information and communication technologies research. *Research Policy*, *28*, 519–543.

Tuttle, B. M., & Kershaw, R. (1998). Information prevention and judgment strategy from a cognitive fit perspective. *Journal of Information Systems*, *12*(1), 1–17.

Typanski, R. E. (1999). Creating an effective information environment. *Information Systems Management*, *16*(2), 32–40.

Vaivio, J. (1999). Exploring a 'non-financial' management accounting change. *Management Accounting Research*, *10*, 409–437.

Venkatraman, N., & Ramanujan, V. (1987). Measuring business economic performance: An examination of method convergence. *Journal of Management*, *13*(1), 109–122.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, *12*(4), 5–34.

White, A. L., Becker, M., & Savage, D. E. (1993). Environmentally smart accounting: Using total cost assessment to advance pollution prevention. *Pollution Prevention Review*, *3*(Summer), 247–259.

Williams, J. J., Macintosh, N. B., & Moore, J. C. (1990). Budget-related behavior in public sector organizations: Some empirical evidence. *Accounting, Organizations and Society*, *15*(3), 221–246.

Wouters, M., Kokke, K., Theeuwes, J., & van Donselaar, K. (1999). Identification of critical operational performance measures – a research note on a benchmarking study in the transportation and distribution sector. *Management Accounting Research*, *10*, 439–452.

Yuthas, K., & Eining, M. M. (1995). An experimental evaluation of measurements of information system effectiveness. *Journal of Information Systems*, *9*(2), 69–84.

Zhang, H. C., Kuo, T. C., Lu, H., & Huang, S. H. (1997). Environmentally conscious design and manufacturing: A state-of-the-art survey. *Journal of Manufacturing Systems*, *16*(5), 352–371.

MANAGING AND CONTROLLING ENVIRONMENTAL PERFORMANCE: EVIDENCE FROM MEXICO

Marc J. Epstein and Priscilla S. Wisner

ABSTRACT

Using data from 236 Mexican manufacturing facilities, we examine the relationship between management control systems and structures and environmental compliance and we test the applicability of management control theory in Mexican industry. We report that success in compliance with environmental regulations is significantly associated with degree of management commitment, planning, belief systems, measurement systems, and rewards. This study contributes to the management control literature by empirically testing the efficacy of management control systems and structures in Mexican industry. It contributes evidence about the implementation of environmental strategies in organizations. Finally, by focusing our analysis on Mexican companies, it gives us a rare view of management control and strategy implementation in a developing economy.

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 115–137

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14005-2

Environmental performance has become a significant dimension in the strategies of many companies, and managers worldwide are focused on identifying and implementing environmental management strategies (Ilinitch, Soderstrom, & Thomas, 1998). In a McKinsey survey of 400 senior corporate executives from around the world, 92% agreed that environmental challenges will be one of the central business issues of the 21st century, and almost all of these executives agreed that their corporations have a responsibility to help control pollution and environmental effects related to their processes and products (Berry & Rondinelli, 1998). The problem is compounded by the daily tradeoffs in decision making that operational managers are faced with, between taking actions that may impact short-term profitability and actions that align with longer-term strategic objectives. Although many senior executives recognize the need for environmental responsiveness, there has generally been a lack of guidance in either the management control or the environmental management literature on actions managers could take to drive an environmental strategy throughout the corporation. Few empirical studies have tested the effectiveness of specific management control systems and structures to push environmental strategy throughout the firm.

This study provides an opportunity to test management accounting and control theories of strategy implementation and provides empirical evidence of the success of various actions that companies take to improve environmental performance. Based on theories developed from the management control and environmental management literature, multiple elements of an organization's management control system are identified. Using data from a cross-section of 236 Mexican manufacturing facilities operating in four key industry groups, we test the association between management control elements and extent of environmental compliance.

This study contributes to the management accounting and control literature by using organizational data to empirically test multiple elements of management control. It extends the study of management control theory and strategy implementation to a developing economy. The study also contributes to the environmental management and strategy literature by identifying actions management can take to more effectively implement an environmental strategy throughout an organization.

The remainder of this paper is organized as follows. In the next section, we review the relevant literature and develop our hypotheses. We then describe our research methods, followed by a discussion of our results and conclusions.

LITERATURE REVIEW

Research into environmental strategy and management is appropriately framed by both contingency theory and the resource-based view of the firm. Contingency theory (Lawrence & Lorsch, 1967) proposes that external pressures are perceived as potential opportunities or possibly threats to the firm, and therefore organizational performance results from the alignment of organizational dynamics with external pressures. Uncertainties in the general business environment increase the likelihood that firms will develop a proactive environmental strategy (Aragon-Correa & Sharma, 2003). Governmental regulations are an externality that are seen as threats by some firms because there are risks associated with non-compliance, and as opportunities by other firms because they can use their compliance to differentiate themselves in the industry and to potentially raise the bar for their competition. Responsiveness to government regulation through management actions is contingent upon management's perception of the degree of threat or opportunity posed by the regulations.

The culture in Mexico prior to the mid-1990s has been described as falling between no effective regulation and one containing some elements of command and control (Nehrt, 1998). The first Mexican environmental law was passed in 1971; however, environmental protection standards were fragmented, government-monitoring resources were low, and enforcement of the existing laws was inconsistent. In the early 1990s, the Mexican government responded to increasing international and domestic pressures for stronger environmental oversight by strengthening the implementation and enforcement of Mexican environmental regulations (Logsdon & Husted, 2000; Rugman & Verbeke, 1998). In 1992, as part of the North American Free Trade Agreement negotiations, Mexico created an environmental oversight agency, PROFEPA (Procuraduria Federal de Proteccion al Ambiente). Environmental oversight was elevated to a secretariat-level agency, SEMARNAP (Secretariat of Environment, Natural Resources, and Fisheries), in 1994. In contrast with the less than 2,000 regulatory inspections conducted between 1971 and 1992, between 1995 and 1998 PROFEPA performed 50,000 industrial inspections (Hufbauer, Esty, Orejas, Rubio, & Schott, 2000). In 1996, the General Law of Ecological Equilibrium and Protection of the Environment was reformed by the Mexican government, strengthening the control of hazardous wastes, increasing criminal penalties, and creating a public database of pollution and violations data (Logsdon & Husted, 2000).

Fundamental to the resource-based view (RBV) of the firm (Barney, 1986, 1991) is that organizations achieve strategic and competitive advantage through the development of resources that are value-creating, rare, and difficult to imitate. Given that the firm is a portfolio of resources (technical, financial, human, etc.) as well as business units and products, competitive advantage and profitability is created by the degree to which management effectively leverages the firm's resources and capabilities (Hamel & Prahalad, 1994). Firms that identify and use management control systems to effectively implement strategy are essentially finding ways to leverage or deploy the assets of a firm in a manner that achieves corporate goals (Simons, 1990, 1994; Dent, 1990). Because an environmental strategy impacts the culture, the operational characteristics, and the longer-term image and reputation of the organization, the outcome is often a set of unique capabilities that are not easily imitated by competitors.

Simons (1987) describes a management control system as the formal, information-based routines and procedures that are used by managers to maintain or alter patterns in an organization's activities. Fisher (1998) describes management control as a framework whereby both exogenous and endogenous factors influence the management control 'package' implemented by the organization. In addition to the formal control mechanisms of goal setting, performance measurement, and incentive compensation systems, management control must also be exerted through organizational structures, culture, human resource management, and other mechanisms.

Often, the study of management control systems has been limited to accounting-based controls and information, which does not reflect the wider range of control mechanisms used by organizations (Langfield-Smith, 1997) and tend to emphasize short-term profitability objectives rather than longer-term strategic objectives (Govindarajan, 1988). Simons' (1994) framework of management control describes four 'levers' of control that help an organization successfully implement strategy: belief systems, boundary systems, diagnostic control systems, and interactive control systems. The balanced scorecard model of Kaplan and Norton (1996, 2000) also describes a comprehensive approach to strategy implementation that links key success factors, key performance indicators, and incentive systems to successfully improving performance. These frameworks resonate with many business managers; however, there have been few empirical tests of these models, as there are very few large-sample studies that enable empirical tests of management control theory (Ittner & Larcker, 1997; Langfield-Smith, 1997; Evans, 1998).

Notable examples include the following. Govindarajan and Gupta (1985) focused on the linkages between type of strategy, incentive systems, and performance at the strategic business unit level in diversified firms, finding that the type of incentive system used interacted with strategy to impact performance. More recently, both Daniel and Reitsperger (1991) and Ittner and Larcker (1997) empirically examined the links between implementing a quality strategy and management control systems. Daniel and Reitsperger found a positive association between pursuit of a zero-defect quality strategy by Japanese firms and the structure of the goal setting and feedback systems used. Ittner and Larcker, using corporate data from four countries, identified management control practices used by firms pursuing a quality strategy, and further linked management control practices and quality performance. Davila (2000) identified factors linking new product development strategy, management control system choices, and performance outcomes.

There has also been very little empirical research into the association between strategy, management control systems, and *environmental* performance (Sharma, 2000; Klassen & McLaughlin, 1996). Management theorists have paid little attention to how corporations can be structured to achieve sustainability and improve environmental performance (Shrivastava, 1995). A focus on the organizational processes, systems, and coordinating mechanisms to implement and support the environmental strategy (Cordano & Frieze, 2000; Sharma & Vredenburg, 1998; Christmann, 2000) is needed. A successful proactive environmental strategy is comprehensive and strategically complex, entailing the coordination of work across functional capabilities and is embedded in the culture of an organization (Buysse & Verbeke, 2004). The effective implementation of environmental strategy requires a well-established, comprehensive system of management control that incorporates strategy, structures, systems, culture, and people (Epstein, 1996). Guided by Simons' (1994) framework of management control and by previous research into effective implementation of environmental management strategies, we have identified seven elements of management control that we believe represent organizational processes.

Plans and Procedures

Planning has been described as one of the most visible and objective parts of the management control system, forming part of the boundary system of the organization and establishing an ex ante form of control (Simons, 1994;

Langfield-Smith, 1997). In Simons' (1994) study of newly appointed top managers, enhanced planning was a key mechanism used to convey strategic agendas and to influence the organization. Davila (2000) found that detailed plans were positively associated with improved new product development performance in his study of 56 companies.

Developing a formal set of procedures and policies or an environmental plan increases self-regulation and is a fundamental part of implementing an effective environmental strategy (Henriques & Sadorsky, 1996). Dasgupta, Hettige, and Wheeler (2000) reported that facilities with an environmental plan exhibit superior environmental performance.[1] Ramus and Steger (2000) found that having a published environmental plan tripled the probability of employees becoming more involved in corporate environmental strategy initiatives. Although there has been some debate about what kind of environmental plan is best, empirical research has found no significant performance effects from using a company-defined environmental plan versus using a certified planning process such as the International Standardization Organization's ISO 14000 or the European Union's Environmental Management and Auditing System (Steger, 2000). What makes a difference is that a company takes a systematic and comprehensive approach to environmental management. The elements that companies include in their environmental plans vary from perhaps only having a mission statement, to having plans and procedures identified to deal with environmental challenges, to incorporating measures and goals into the planning process. Consistent with Simons' (1994), we hypothesize that facilities that have more comprehensive sets of plans that include multiple planning elements will be more likely to reach compliance goals than facilities that do not have plans that are as comprehensive.

H1. Having a more comprehensive environmental plan will be positively associated with environmental compliance.

Belief Systems

Management control is exerted through having a shared belief system or a common set of core values, as this helps to align the decision making of the employees with the mission of the organization (Simons, 1994; Merchant, 1985). Organizational culture is a form of social control, helping to promote and ensure goal congruence, and is thought to be especially effective when uncertainties are high (Flamholtz, Das, & Tsui, 1985). When employees see that their values are reflected in the goals of the company, they become more

committed to their jobs and to the company (Packard & Reinhardt, 2000; Ramus & Steger, 2000). Having a strong sense of internal mission increases employee commitment to the organization, and also lessens uncertainties about appropriate courses of action to take. We therefore hypothesize that:

H2. The strength of the belief system will be positively associated with environmental compliance.

Management Commitment

Interactive control systems are those that managers use to personally involve themselves in the decision-making processes of the organization (Simons, 1994). The degree of management attention or commitment signals to the employees the importance of a strategic initiative, and is especially important when there exist many strategic uncertainties. Top management attention and commitment is necessary to support many of the strategic initiatives required to implement an effective environmental strategy, including initiatives that create changes in corporate culture, marketing strategies, and operational approaches and capabilities (Kolk, 2000; Cordano & Frieze, 2000; Menon & Menon, 1997). The degree of management commitment not only signals the tone at the top, but it also gives internal credibility to the environmental initiatives that must be undertaken to improve performance (Epstein & Roy, 1998). Judge and Douglas (1998) found a positive environmental performance impact when companies explicitly integrated environmental issues into the strategic planning process. We hypothesize that environmental compliance will be better in companies that have strong management commitment to an environmental strategy, because this will help to focus manager and worker attention toward making decisions that improve environmental compliance.

H3. The degree of management commitment to environmental issues will impact environmental compliance.

Reporting Structure

Senior managers have to make daily choices about how to allocate their limited amount of time and attention. Top management, by necessity, must delegate decision-making and control responsibilities to subordinates in the organization. The reporting structure in an organization is both a signal about the importance of the environmental strategy and an operational mechanism to strengthen environmental management (Epstein, 1996). One

critical element of an environmental management system is that the environmental manager must have influence and authority over decision making (Cordano & Frieze, 2000). To test the influence of reporting systems on environmental compliance, the following hypothesis is proposed.

H4. Environmental compliance will be stronger in firms where the environmental manager reports directly to top management.

Integration

Decentralized decision-making is one response to increased complexity (Simons, 1990). To control complex processes, employee learning, involvement, and accountability must be pushed downward and throughout the organization. Kaplan and Norton (2000) refer to this as the movement of strategy from the 10 (the senior executive team) to the 10,000 (everyone in the company). Although individuals possess the expertise and knowledge needed to adjust to changes in the corporate environment, as complexity increases no one individual or unit will possess all the knowledge or capabilities within the organization. Also, as complexity increases, more exceptions to decision making occurs, potentially overloading an organization's hierarchy (Govindarajan, 1988). The complexity of corporate environmentalism has increased dramatically in the past decades. Externally, pressures have increased from a diverse set of stakeholders, including governments, customers, investors, environmental groups, and communities. Environmental management practices therefore involve strategic, marketing, financial, human resources, legal, operational, and product development functions. Successfully implementing an environmental strategy is often improved by integrating responsibility for environmental performance to employees throughout the operational and support functions of the organization (Shrivastava, 1995; Aragon-Correa, 1998; Sharma & Vredenburg, 1998). Positive impacts on environmental performance have been associated with environmental training (Dasgupta et al., 2000) and increased functional responsibility (Judge & Douglas, 1998). This suggests that the extent to which environmental responsibilities are integrated throughout the organization will impact corporate environmental compliance, which leads us to propose the following hypothesis.

H5. The degree of integration of environmental responsibilities will be positively associated with environmental compliance.

Measurement System

The intensity of monitoring results is a control system attribute (Simons, 1987). Companies create value through selecting performance measures that link to strategic objectives, that align with key performance criteria, and that measure the results of processes that the company is trying to manage (Merchant, 1998; Ittner & Larcker, 1998). In environmental management, government regulations typically specify types of discharges to be measured, frequency of measurement, and allowable error tolerances. However, the actual measurement practices of firms vary widely. Some firms do not measure discharges or implement a weaker measurement system than is required by regulation. Other firms only measure to ensure that they are meeting regulatory requirements. Still other firms choose to implement a stronger measurement system that exceeds regulatory requirement. One of the critical choices in environmental measurement systems is the strength of the pollution discharge measurement system (Berry & Rondinelli, 1998). The measurement system of a company is both an ex ante and an ex post control system. Performance measures serve as an ex ante control by specifying expected performance outcomes, but are primarily seen as ex post diagnostic control tools that give information about whether certain goals and targets have been met (Simons, 1994; Flamholtz et al., 1985). Therefore, we expect the strength of the discharge measurement system to be an ex ante signal to the organization about the importance of environmental compliance, and also to serve as an ex post control to ensure that compliance standards are being met.

H6. Stronger discharge measurement systems will be positively associated with environmental compliance.

Reward System

Performance-dependent rewards are used as part of a management control strategy in two important ways (Merchant, 1998). First, they provide information to employees about the relative importance of often-competing objectives, helping employees to focus their efforts. Second, they motivate employees to work toward a defined goal in order to achieve the rewards. Organizational performance is higher when reward systems are matched to business unit strategy (Govindarajan & Gupta, 1985). Bonus or incentive compensation is one management control tool that aligns the interests of those who define the strategy with the interests of those who execute the

strategy (Fisher & Govindarajan, 1993; Simons, 1987). In environmental management research, Ramus and Steger (2000) found a positive link between rewards and recognition and employee involvement in environmental initiatives, but did not test the impact of this link on corporate environmental performance. We test whether the reward system also influences environmental compliance outcomes with the following hypothesis.

H7. Rewarding employees for environmental achievements will be positively associated with environmental compliance.

METHODS

In late 1995, researchers from the World Bank, along with various Mexican government and academic partners, conducted an extensive study into corporate environmental practices in Mexico. Supported by the Mexican Secretariat of Environment, Natural Resources, and Fisheries (*Secretaria de Medio Ambiente, Recursos Naturales y Pesca – SEMARNAP*) and the Mexican Association of Industries, a team of World Bank researchers developed a detailed questionnaire focusing on organizational structure, environmental challenges, and environmental performance issues. Four industry groups were chosen to be the focus of the study: food, chemicals, non-metallic minerals, and metals. These industries were estimated to account for 75–95% of Mexico's total industrial pollution. Within these industries, 236 firms that were located in the industrial corridors of Monterey, Guadalajara, and Mexico City participated in the study. The size of the facilities participating in the study ranged from small (<100 employees) to large (>250 employees). Table 1 provides size and sector information for the manufacturing facilities.

A research team from the Monterrey Institute of Technology (*Instituto Technologico y de Estudios Superiores de Monterrey – ITESM*) visited each of the facilities and conducted structured interviews in Spanish.[2] The scope of interview data included firm characteristics, sales and marketing information, organizational systems and structures, environmental compliance data, and other operational data. The survey instruments were completed by the research team through interviews with various management and operational personnel in each plant. All of the data were provided by plant personnel; however, the research team corroborated some of the data as part of their plant visits. Owing to strict confidentiality agreements between the

Table 1. Size and Sector Data for Sample Facilities.

	Food	Chemicals	Non-Metallic Minerals	Metals	Total
Small (16–100 employees)	19	23	27	23	92 (39%)
Medium (100–250 employees)	22	21	12	18	73 (31%)
Large (over 250 employees)	21	18	12	20	71 (30%)
Total	62 (26%)	62 (26%)	51 (22%)	61 (26%)	236 (100%)

research team and the plants participating in the study, no identifying characteristics about the plants or the personnel interviewed were reported.

Variables

The variables used in the data analysis were generated directly from the responses received on the interview instrument. The variables are described in Table 2; summary statistics are presented in Table 3.

A number of potential control variables were evaluated: size, age of plant, percentage of exports, and regional destination of exports. Only size was positively associated with performance ($p < 0.0001$), consistent with previous research reporting firm size as a moderating variable on environmental performance (Christmann, 2000). The Kruskal–Wallis test, commonly used for nonparametric data, indicated that the facilities in the non-metallic minerals industry group were significantly smaller than the other industry groups. We have therefore included size as a control variable.

The dependent variable, *compliance*, was a self-report of environmental compliance by the facility. The assurance of confidentiality to the responding facilities prevents us from triangulating the self-report of compliance with externally reported data. However, the self-report of compliance was validated independently by the ITESM research team at each of the research locations; the ITESM team reported strong correlation between the self-report score of environmental compliance and observed conditions (Dasgupta et al., 2000). Also, we triangulated the compliance data with two additional pieces of facility data. The first was a listing of environmental *performance actions* undertaken by the facility, such as reducing or eliminating toxic materials, installing treatment equipment, waste reduction or

Table 2. Variables, Survey Questions, and Responses.

Plans and procedures

Based on the ISO 14000 international environmental certification standard, respondents were asked about the use of eight planning and procedural elements in their facilities: a formal environmental management policy with written procedures, mission statement, plans and procedures for emergencies, community interaction plans, procedures beyond compliance, measures, goals, and waste reduction plans. The data were clustered into three evenly distributed groups: low, medium, and high planning.[a]

1 – Low planning (0–2 elements)
2 – Medium planning (3–5 elements)
3 – High planning (6–8 elements)

Belief systems

How much does the internal mission of the company influence the environmental actions taken by this company?

1 – It has little influence.
2 – It has some influence.
3 – It has a definite influence.
4 – It has a very definite influence.

Management commitment

Which of the following policies describe the emphasis that upper management places on environmental matters?

1 – Our management does not motivate us to comply with the law.
2 – Our management wants us to comply with the law if and only if there are risks of penalties.
3 – Management is interested in complying with the law, but is not interested in going further.
4 – It is a topic of high importance.
5 – It is a topic of highest priority for our management.

Reporting structure

To whom does the person responsible for environmental matters in this plant report?

0 – Functional manager (e.g., operations, human resources, legal, security).
1 – Top management (e.g., owner, board of directors, general or regional manager).

Integration

In how many operational and support functions of the plant are there employees with environmental responsibilities? (0 – None; 1 – One; 2 – More than one)

Measurement system

How do you measure the environmental discharges from your plant?

0 – Discharges not measured.
1 – Discharges are measured.
2 – Discharges are measured in accordance with environmental regulations.
3 – Discharges are measured in ways that exceed environmental regulations.

Table 2. (*Continued*)

Reward system
Do you financially or otherwise reward employees for improving the environmental performance of this company? (0 – No; 1 – Yes)
Compliance
What is your best characterization about how you would describe your plant's environmental performance?
1 – It is rare that we comply with MX regulation.
2 – We normally do not comply with MX regulations.
3 – We normally comply with MX regulations; however, sometimes we miss in specific areas.
4 – We consistently comply with MX environmental regulations.
5 – We exceed the required regulations and have a world-class environmental program.

[a]The data were clustered into three response categories to mitigate problems with small cell sizes using the raw data.

recycling actions, and process changes. The second was a report of *environmental outcomes*, defined as more efficient use of energy, materials, or water, as well as economic improvements due to pollution prevention. We would expect to find a positive association between *performance actions* and *compliance*, and between *compliance* and *environmental outcomes*, as companies that take proactive actions to impact performance would be more likely to comply with regulations, and facilities with better environmental compliance would be more likely to use resources more efficiently. Both performance actions and environmental outcomes were significantly correlated with compliance (0.26 and 0.40, respectively). Further, a χ^2-test of independence (size-controlled) between each of these measures and compliance showed significant association between *performance actions* and *compliance* ($\chi^2 = 8.76$, $p < 0.0031$) and between *compliance* and *environmental outcomes* ($\chi^2 = 26.20$, $p < 0.0001$). These validity checks strengthen our confidence in the self-report measure of *compliance*.

Model

The questionnaires used in the study were designed such that the levels of the response are equi-spaced in intent. Agresti (1990) states that the scores should reflect the insights about the way in which the classification was constructed and used. We therefore employ a technique described in Agresti (1990) to extend loglinear logit modeling to account for ordinality of the predictors and assign the row and column scores as equi-spaced. The

Table 3. Descriptive Data: Variable Ranges, Means, Spearman Correlations (Two-Tailed) and Number of Respondents.

		range	mean	1	2	3	4	5	6	7	8
1	Plans and procedures	1–3	1.99	1.00							
				186							
2	Belief system	1–4	2.84	0.23*	1.00						
				100	135						
3	Management commitment	1–5	3.78	0.50***	0.41***	1.00					
				185	131	227					
4	Reporting system	0–1	0.67	−0.09	−0.07	−0.11	1.00				
				167	117	202	208				
5	Integration	0–2	0.89	0.35***	0.13	0.29***	−0.11	1.00			
				186	135	227	208	236			
6	Measurement system	0–3	1.60	0.53	0.19*	0.48***	−0.18*	0.25***	1.00		
				180***	131	218	201	225	235		
7	Reward system	0–1	0.19	0.25***	0.13	0.24***	−0.21**	0.12	0.22***	1.00	
				184	134	225	206	233	223	233	
8	Compliance	1–5	3.46	0.43***	0.28**	0.52***	−0.13	0.14*	0.36***	0.20**	1.00
				184	128	220	199	224	215	222	224

Significance levels:
***$p < 0.001$;
**$p < 0.01$;
*$p < 0.05$.

commonly employed model in this case is further described in Nelder and Wedderburn (1972) as a Generalized Linear Model with a log-link function and assumed Poisson distributed counts. We begin with a row by column contingency table and assumed Poisson distributed counts in the cells. Common loglinear models fail to recognize departures from independence that reflect stochastic ordering in the predictors of interest; we therefore include a term in the model to describe the order-based association between the rows and columns. The *size-controlled* model is

$$\log\, m_{ijk} = \mu + \alpha_i + \tau_j + \delta_k + \beta u_i v_j + \alpha\delta_{ik} + \tau\delta_{jk}$$

where m_{ijk} the count associated with the ijkth cell for ith row, jth column and kth size ($i = 1, \ldots, I$, $j = 1, \ldots, J$ and $k = 1, \ldots, K$), u_i the ith fixed column order score, v_j the jth fixed row order score, μ the overall mean, α_i the ith row effect, τ_j the jth column effect, δ_k the kth size effect, β the association, $\alpha\delta_{ik}$ the ith row by kth size interaction, $\tau\delta_{jk}$ the jth column by kth size interaction.

We evaluated the outcome of each hypothesis test using the beta term, which describes a measure of linear-by-linear association between the predictors of interest. All parameters in the model were estimated using a ridge-stabilizing Newton–Raphson, an iterative algorithm that maximizes the log-likelihood function with respect to the model. We further verified the validity of each model by examining the standardized Spearman residuals and plots of observed versus fitted responses; no deviations were found.

RESULTS AND DISCUSSION

Table 4 contains a summary of the data outcomes and the statistical results of our analyses.

The results confirm our first hypothesis (H1), finding a positive and significant relationship between planning and environmental compliance. Looking at the planning response ranges, we see that environmental compliance improves from a mean score of 2.25 (less than compliance; see Table 2 for a breakdown of the environmental compliance ranges) when planning is low, to a mean score of 3.05 (just above compliance) when planning is high in a manufacturing facility. It is clear from the pattern of the data and the statistical outcome that facilities that create more comprehensive plans for environmental compliance also enjoy better compliance outcomes.

Table 4. Statistical Results.

Management Control Elements[a] (Independent Variables)	Compliance (DV)	df	Estimate	Standard Error	χ^2	*p*-value
Plans and procedures ($n = 184$)		1	0.7230	0.1626	19.78	<0.0001
1. Low	2.25					
2. Medium	2.52					
3. High	3.05					
Belief system ($n = 128$)		1	0.3557	0.1130	9.90	0.0016
1. Little influence	2.14					
2. Some influence	1.87					
3. Definite influence	2.30					
4. Very definite influence	2.50					
Mgt. Commitment ($n = 220$)		1	0.9334	0.1474	40.11	0.0001
1. Lowest	0.50					
2. Some	1.50					
3. Medium	2.08					
4. High	2.61					
5. Highest	3.07					
Reporting structure ($n = 199$)		1	-0.1808	0.2021	0.80	0.3708
0. Functional mgt.	2.66					
1. Top management	2.42					
Integration ($n = 224$)		1	0.1531	0.1001	2.34	0.1262
1. None	2.37					
2. One	2.34					
3. $>$one	2.69					
Measurement system ($n = 215$)		1	0.3651	0.0819	19.87	<0.0001
0. Not measured	2.06					
1. Measured	2.75					
2. In accordance with regs.	3.05					
3. Exceeds regs.	3.17					
Reward system ($n = 222$)		1	0.5212	0.2347	4.93	0.0264
1. No	2.19					
2. Yes	2.62					

[a]The *n* of responses differs because this represents the number of facilities who provided both the management control element of interest *and* compliance data. Not all facilities responded to all questions.

Hypothesis H2 was also strongly supported by the analysis, as the relationship between the belief system and environmental compliance was positive and significant. Although there is an unexpected small drop in environmental compliance as the degree of influence increases from 'little' to 'some', we can clearly see that the general pattern of improvement is as expected (the greater the influence, the better the environmental compliance outcome). This result provides evidence for the theory of Merchant (1985) and Simons (1994) that having a shared belief system or core values influences compliance outcomes. It also provides evidence to support Govindarajan and Fisher's (1990) contention that socialization control may be an important factor in non-U.S. cultures.

The statistical relationship between management commitment and environmental compliance (H3) was powerful. The results of the χ^2-test demonstrate a strong relationship, and by examining the mean environmental compliance value for each level of management commitment, we see how much this control element drives compliance. For facilities reporting the lowest level of management commitment, environmental compliance was abysmal. For each subsequent increase in management commitment to environmental compliance, we can see that performance was substantially improved. This tells us that environmental compliance is highly sensitive to degree of management commitment, and is driven from the top down through the organization. We believe that the influence of senior management commitment would be especially high when strategic uncertainties are also high, which is perhaps one explanation as to why this variable shows such strong results in this dataset. During the 1990s in Mexico, environmental regulations were changing due to influence from the North American Free Trade Agreement and also due to increased public demands for corporate environmental responsibility. The changes would have put a new and different kind of pressure on companies to meet environmental performance standards, and would have increased the degree of uncertainty in corporate environmental planning. The changes would also require strong management commitment to respond to the increasing pressures of the Mexican regulatory environment. This may be one reason why management commitment has such a striking influence on environmental performance in Mexico.

Hypothesis H4, that compliance would be better when the environmental manager reports to top management as opposed to functional management, was not supported by the data. We further analyzed the data to determine if there was an interaction with the size variable; however, the reported results were consistent across size of facility. Our results did not support our

expectation that having the environmental manager report to top management would result in improved compliance.

Hypothesis H5, proposed that environmental compliance is better in organizations that integrate environmental accountability and responsibility throughout the organization. The integration hypothesis was tested by comparing environmental compliance in plants according to the number of operational and functional areas with environmental responsibilities. The data did not support our hypothesis. However, as we see in the environmental compliance means related to the degree of integration, compliance was not significantly improved as functional coverage increased from 'none' to 'one' operational or support area, but improvement is triggered when employees in more than one functional area have environmental management responsibilities. In companies reporting functional responsibility for environmental compliance, operations was the function most frequently cited by the respondents ($n = 107$), followed by production design ($n = 47$), process design ($n = 40$), and engineering ($n = 44$). These responses, along with the data showing that environmental compliance is improved when more than one functional area is involved, suggest that operational responsibility alone is not enough to influence environmental compliance. This suggests that to effectively impact environmental outcomes, firms need to involve functional areas beyond operations. Functions such as product design, process design, and engineering would enable firms to more proactively exert control over environmental outcomes, as opposed to reacting to environmental problems after they have occurred.

The positive relationship between strength of the measurement system and environmental compliance was strongly supported in the test of hypothesis H6. In the case of environmental compliance, measurement of facility discharges is directly related to pollution control, which is a key performance variable for government regulators. The data patterns show that there is a substantial increase in compliance among firms who measure discharges as opposed to no measurement, and there is additional improvement as the strength of the measurement system increases.

Hypothesis H7, the association of a reward system and environmental compliance, was positive and significant. By looking more closely at the reward data, we are able to gain some interesting insights. Of the facilities reporting a reward system for environmental compliance, environmental personnel were eligible for the awards in 34% of the facilities, non-environmental managers were eligible in 16% of the facilities, and non -environmental workers were eligible in 82% of the facilities. Most of the earlier empirical studies on the use of reward systems to control

performance have focused on the compensation structure of the business unit manager (Govindarajan & Gupta, 1985; Fisher & Govindarajan, 1993; Simons, 1987). Our data demonstrate that a reward system that primarily impacts *non-managerial workers* is also an effective means to control performance outcomes. Another unique and interesting aspect of the reward data is that the reward structure is varied between monetary rewards (45% of facilities), 'in kind' awards (34%), recognition (55%), and other (5%). These data indicate that achieving the compliance outcome is positively linked with the broad concept of being rewarded, and not just receiving a monetary reward.

CONCLUSION

Our analysis of data from a cross-section of industries in Mexico reports results that are consistent with the theoretical frameworks of strategy implementation. We report that stronger environmental compliance is significantly associated with management commitment, planning, belief systems, measurement systems, and rewards.

Although the structure of the data enables us to measure the association between the management control elements and environmental compliance, the data is cross-sectional and snapshot in nature, therefore causality cannot be inferred. However, using insights gained from management control and environmental management literature, we are able to hypothesize, test, and confirm the direction of the relationships. Another limitation is that the data were self-report measures, which are subject to perceptual and bias errors. The anonymity of the respondents may have minimized response bias, but also restricted further clarification or external triangulation of responses. Two factors strengthen confidence in the reliability of the self-report measures. First, the ITESM research team partially verified responses through on-site observations and questioning. Also, we statistically triangulated our dependent variable, environmental compliance, across multiple measures. Finally, there is a potential bias in the data collection if a single interviewee provides the data for the dependent and independent variables. The interviewers reported that multiple managers were interviewed in the plant visits, and that the interviewers also corroborated some of the data. However, because no information was reported on who responded to what questions in the survey, it is impossible to test for rater-driven biases in the data.

Potential extensions of this research stream include analyses of management control frameworks used by companies that operate in more than one

country. Understanding the interactions between culture and management control would be a further extension of the framework proposed by Simons (1994) and would provide valuable insights for managers of multi-national companies. It would also be worthwhile to evaluate if the choice of management control elements is dependent upon the external pressures on the company. For example, Evans (1998) suggests that the role of management control will change in the health care industry as government regulation changes. Government regulation is also a strong influence on a company's environmental responsiveness; therefore, the role of the MCS might be different in economies with different governmental pressures, or as government regulation changes in a single economy. Another potential extension would be to use hierarchical analysis to examine the effects of these management control elements within each industry, or the impacts of other interactions within the data. Effectively evaluating these interaction effects would require sufficient data points to support the analysis.

The findings of this study have both academic and managerial relevance. This study contributes to the management control literature by empirically linking the implementation of strategy to compliance outcomes through a variety of control elements. This study extends academic research of management control systems into companies operating in a developing economy, with the results indicating that the theoretical constructs of management control significantly influence outcomes in the developing economy of Mexico. By identifying control elements that are effective in managing environmental compliance, we also contribute to the environmental management literature. The results of this study also provide guidance to managers about how to effectively drive a strategy through an organization. This study focused on the implementation of an environmental strategy. However, the results of planning, having a strong belief system, management commitment, integration, measurements, and reward systems provide insights to managers implementing other strategies as well, for example quality, innovation, or e-commerce strategies.

NOTES

1. Dasgupta et al. (2000) helped to develop the data collection protocol for this study and analyzed the relationship between various internal and external influences and whether or not facilities comply with Mexican environmental regulations. Using the same data set, we extend Dasgupta et al. by testing hypotheses specifically related to management control theory. Our study also differs in our choice of independent variables, methodology, and the dependent variable.

2. The complete survey instrument (in Spanish) is available on the World Bank website. An English translation is available from the authors.

REFERENCES

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Aragon-Correa, J. A. (1998). Strategic proactivity and firm approach to the natural environment. *Academy of Management Journal*, *41*(5), 556–567.

Aragon-Correa, J. A., & Sharma, S. (2003). A contingent resource-based view of proactive corporate environmental strategy. *Academy of Management Review*, *28*, 71–88.

Barney, J. B. (1986). Strategic market factors: Expectations, luck, and business strategy. *Management Science*, *32*, 1231–1241.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*, 99–120.

Berry, M. A., & Rondinelli, D. A. (1998). Proactive corporate environment management: A new industrial revolution. *The Academy of Management Executive*, *12*(2), 38–50.

Buysse, K., & Verbeke, A. (2004). Environmental strategy choice and financial profitability, In: S. M. Lundan (Ed.), *Multinationals environment and global competition* (pp. 43–63). Oxford: Elsevier.

Christmann, P. (2000). Effects of "best practices" of environmental management on cost advantage: The role of complementary assets. *Academy of Management Journal*, *43*(4), 663–680.

Cordano, M., & Frieze, I. H. (2000). Pollution reduction preferences of U.S. environmental managers: Applying Ajzen's theory of planned behavior. *Academy of Management Journal*, *43*(4), 627–641.

Daniel, S. J., & Reitsperger, W. D. (1991). Linking quality strategy with management control systems: Empirical evidence from Japanese industry. *Accounting, Organizations and Society*, *16*(7), 601–618.

Dasgupta, S., Hettige, M., & Wheeler, D. (2000). What improves environmental compliance? Evidence from Mexico. *Journal of Environmental Economics and Management*, *39*(1), 39–66.

Davila, T. (2000). An empirical study on the drivers of management control systems' design in new product development. *Accounting, Organizations and Society*, *25*, 383–409.

Dent, J. F. (1990). Strategy, organization and control: Some possibilities for accounting research. *Accounting, Organizations and Society*, *15*(1–2), 3–25.

Epstein, M. J. (1996). *Measuring corporate environmental performance*. Chicago: Irwin Professional Publishing.

Epstein, M. J., & Roy, M. J. (1998). Managing corporate environmental performance: A multinational perspective. *European Management Journal*, *16*(3), 284–296.

Evans, J. H. (1998). Cost management and management control in health care organizations: Recent opportunities. *Behavioral Research in Accounting*, *10*(Suppl), 78–93.

Fisher, J. (1998). Contingency theory, management control systems and firm outcomes: Past results and future directions. *Behavioral Research in Accounting*, *10*, 47–64.

Fisher, J., & Govindarajan, V. (1993). Incentive compensation design, strategic business unit mission, and competitive strategy. *Journal of Management Accounting Research*, *5*, 129–145.

Flamholtz, E. G., Das, T. K., & Tsui, A. S. (1985). Toward an integrative framework of organizational control. *Accounting, Organizations and Society*, *10*(1), 35–50.

Govindarajan, V. (1988). A contingency approach to strategy implementation at the business-unit level: Integrating administrative mechanisms with strategy. *Academy of Management Journal*, *31*(4), 828–853.

Govindarajan, V., & Fisher, J. (1990). Strategy control, systems, and resource sharing: Effects on business-unit performance. *Academy of Management Journal*, *33*(2), 259–285.

Govindarajan, V., & Gupta, A. K. (1985). Linking control systems to business unit strategy: Impact on performance. *Accounting, Organizations and Society*, *10*(1), 51–66.

Hamel, G., & Prahalad, C. K. (1994). *Competing for the future*. Boston: Harvard Business School Press.

Henriques, I., & Sadorsky, P. (1996). The determinants of an environmentally responsive firm: An empirical approach. *Journal of Environmental Economics and Management*, *30*, 381–395.

Hufbauer, G. C., Esty, D. C., Orejas, D., Rubio, L., & Schott, J. J. (2000). *NAFTA and the environment: Seven years later*. Washington, DC: Institute for International Economics.

Ilinitch, A. Y., Soderstrom, N. S., & Thomas, T. E. (1998). Measuring corporate environmental performance. *Journal of Accounting and Public Policy*, *17*, 383–408.

Ittner, C. P., & Larcker, D. F. (1997). Quality strategy, strategic control, systems and organizational performance. *Accounting, Organizations and Society*, *22*(3–4), 293–314.

Ittner, C. P., & Larcker, D. F. (1998). Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research*, *10*, 205–238.

Judge, W. Q., Jr., & Douglas, T. D. (1998). Performance implications of incorporating natural environmental issues into the strategic planning process: An empirical assessment. *Journal of Management Studies*, *35*(2), 241–262.

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: Translating strategy into action*. Boston: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. (2000). *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Boston: Harvard Business School Press.

Klassen, R. S., & McLaughlin, C. P. (1996). The impact of environmental performance on firm performance. *Management Science*, *42*(8), 1199–1214.

Kolk, A. (2000). *Economics of environmental management*. Harlow, England: Pearson Education Limited.

Langfield-Smith, K. (1997). Management control systems and strategy: A critical review. *Accounting, Organizations and Society*, *22*(2), 207–232.

Lawrence, P. R., & Lorsch, J. W. (1967). *Organization and environment: Managing differentiation and integration*. Boston: Harvard Business School Press.

Logsdon, J. M., & Husted, B. W. (2000). Mexico's environmental performance under NAFTA: The first 5 years. *Journal of Environment & Development*, *9*(4), 370–383.

Menon, A., & Menon, A. (1997). Enviropreneurial marketing strategy: The emergence of corporate environmentalism as market strategy. *Journal of Marketing*, *61*(1), 51–67.

Merchant, K. A. (1985). Organizational controls and discretionary program decision making: A field study. *Accounting Organizations and Society*, *10*(1), 67–85.

Merchant, K. A. (1998). *Modern management control systems*. Upper Saddle River, NJ: Prentice-Hall.

Nehrt, C. (1998). Maintainability of first mover advantages when environmental regulations differ between countries. *Academy of Management Review*, *23*(1), 77–97.

Nelder, J., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of Royal Statistical Society: Series A (Statistics in Society)*, *135*, 370–384.

Packard, K. O., & Reinhardt, F. (2000). What every executive needs to know about global warming. *Harvard Business Review,* (July–August), 129–135.

PROFEPA, *El Auditoria Ambiental en Mexico* (The Environmental Audit in Mexico). Procuraduria Federal de Proteccion al Ambiente 2000.

Ramus, C. A., & Steger, U. (2000). The roles of supervisory support behaviors and environmental policy in employee "ecoinitiatives" at leading-edge European companies. *Academy of Management Journal*, *43*(4), 605–626.

Rugman, A. M., & Verbeke, A. (1998). Corporate strategy and international environmental policy. *Journal of International Business Studies*, *29*(4), 819–833.

Sharma, S. (2000). Managerial interpretations and organizational context as predictors of corporate choice of environmental strategy. *Academy of Management Journal*, *43*(4), 681–697.

Sharma, S., & Vredenburg, H. (1998). Proactive corporate environmental strategy and the development of competitively valuable organizational capabilities. *Strategic Management Journal*, *19*, 729–753.

Shrivastava, P. (1995). The role of corporations in achieving ecological sustainability. *Academy of Management Review*, *20*(4), 936–960.

Simons, R. (1987). Accounting control systems and business strategy: An empirical analysis. *Accounting Organizations and Society*, *12*(4), 357–374.

Simons, R. (1990). The role of management control systems in creating competitive advantage: New perspectives. *Accounting Organizations and Society*, *15*(1–2), 127–143.

Simons, R. (1994). How top managers use control systems as levers of strategic renewal. *Strategic Management Journal*, *15*, 169–189.

Steger, U. (2000). Environmental management systems: Empirical evidence and further perspectives. *European Management Journal*, *18*(1), 23–37.

This page intentionally left blank

STRATEGIC ORGANIZATIONAL DEVELOPMENT AND FINANCIAL PERFORMANCE: IMPLICATIONS FOR ACCOUNTING, INFORMATION, AND CONTROL

Eric G. Flamholtz

ABSTRACT

This paper examines the implications for accounting, information, and control of a growing body of research to develop and empirically test of a holistic model of organizational success and failure in entrepreneurial organizations at different stages of growth. It builds upon previous work by Falmholtz and colleagues on developing a model of organizational success and failure. It also builds upon a perspective previously developed by Flamholtz, which presents a broader view of the role of accounting control systems in an organizational context.

The initial model proposes that there are six key factors or ''strategic building blocks'' of successful organizations, and the six key variables must be designed as a holistic system, which has been termed ''The Pyramid of Organizational Development''. The model together with the growing body of research designed to assess its validity has significant implications for accounting, information, and control.

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 139–165

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14006-4

BACKGROUND

In recent years, most industries throughout the world have witnessed successes and failures of seemingly similar companies. Organizations such as Microsoft, Southwest Airlines, Nike, and Wal-Mart become dominant forces in their industries while other comparable organizations such as Apple Computer, People Express, LA-Gear, and K-Mart have experienced difficulties and decline after a period of promising initial growth (Flamholtz & Randle, 1998).

The result is an increased need for a better understanding of the management of organizational growth and the determinants of success and failure over the long term. More specifically, why do some organizations continue to be successful over the long term while others, with equally promising starts, experience difficulties and even failure?

To help answer this question, Flamholtz (1995) presented a framework entitled the "Pyramid of Organizational Development" that identified six key "strategic building blocks" of successful organizations. Subsequently, Flamholtz and colleagues have engaged in a program of empirical research to assess the validity of the model and various hypotheses and implications derived from it.

The next section provides a review of the key aspects of the framework relevant to this research. The third section will survey the empirical research, which has been conducted to date to assess the validity of various hypotheses derived from the framework. Finally, the implications of this research for theory and practice will be considered in the final section.

THE THEORETICAL FRAMEWORK

The framework consists of four key parts: (1) a "strategic organizational development" model, (2) a life cycle model, (3) a model of the levels of strategic organizational development required at each stage of growth, and (4) a framework for the dysfunctional consequences, which occur when suboptimal strategic organizational development occurs. These are described, in turn, below.

The Model for Strategic Organizational Development

The initial premise or hypothesis underlying this framework is that organizations must perform certain tasks to be successful at each stage of their

growth. The six key tasks of strategic organizational, all of which have been supported by previous research are:

- Identification and definition of a viable market niche (Aldrich, 1979; Brittain & Freeman, 1980; Freeman & Hannan, 1983)
- Development of products or services for the chosen market niche (Burns & Stalker, 1961; Midgley, 1981)
- Acquisition and development of resources required to operate the firm (Pfeffer & Salancik, 1978; Brittain & Freeman, 1980; Carroll & Yangchung, 1986)
- Development of day-to-day operational systems (Starbuck, 1965).
- Development of the management systems necessary for the long-term functioning of the organization (Child & Keiser, 1981; Tushman, Virany, & Romanelli, 1985).
- Development of the organizational culture that management feels necessary to guide the firm (Peters & Waterman, 1982; Walton, 1986).

Each of these key tasks will be discussed in detail below.

Identification of Market Segment and Niche

The first challenge for a new venture in organizational survival or success is to identify a market need for a marketable service or product. The chances of organizational success are enhanced to the extent that the firm is successful in this step (Flamholtz, 1995).

The challenge is not merely in identifying the market but also, if possible, to capture a ''market niche,'' a relatively protected place that would give the company sustainable competitive advantages. Failing to define a niche or mistakenly abandoning the historical niche can cause an organization to experience difficulties and even failure. The process of identifying the market involves the development of a strategic market plan to identify potential customers and their needs and the creation of a competitive strategy (Flamholtz, 1995).

Development of Products and Services

The second challenge or strategic building block involves the development of products and/or services. This process can also be called ''productization,'' which refers to the process of analyzing the needs of customers in the target market, designing the product and developing the ability to produce it (Flamholtz & Randle, 2000). For a production firm, this stage involves the design and manufacturing phases, whereas for a service firm, this stage

involves forming a system for providing services to the customers (Flamholtz & Randle, 2000).

The success of this stage is highly related to the previous critical task, proper definition of the market niche (Flamholtz, 1995). Unless a firm fully understands the needs of the market, it cannot satisfy those needs in productization.

Acquiring Resources

Success in identifying a market niche and productization will create increased demand for a firm's products or services. Consequently, the resources of the firm will be spread very thin (Flamholtz, 1995). The organization will require additional physical, financial, and human resources. This is the point at which the entrepreneur/s should start thinking about the long-term vitality of the firm and procure all the necessary resources to survive the pressure of current and future increase in demands (Flamholtz & Randle, 2000).

Development of Operational Systems

The fourth critical task is the development of basic day-to-day operational systems, which include accounting, billing, collection, advertising, personnel recruiting and training, sales, production, delivery, and related systems (Flamholtz, 1995). Entrepreneurial companies tend to quickly outgrow the administrative systems available to operate them. Therefore, it is necessary to develop sufficient operational systems, on time, to build a successful organization. In contrast, large established companies might have developed overly complicated operational systems. In this case, the success of the organization depends on the reengineering of operational systems (Flamholtz, 1995).

Development of Management Systems

The fifth step is to develop the management systems, which is essential for the long-term viability of the firm (Flamholtz & Randle, 2000). Management systems include systems for planning, organization, management development, and control. Planning systems involve planning for the overall development of the organization and the development of scheduling and budgeting operations. It includes strategic planning, operational planning, and contingency planning (Flamholtz, 1995). The mere existence of planning activities does not indicate that the firm has a planning system. A planning system ensures that planning activities are strategic and ongoing.

Organizational structure involves the ways in which people are organized and activities are coordinated. As with the planning activities, success depends not on the mere existence of a structure, but on the match between the structure and business strategy (Flamholtz, 1995).

The process of planned development of the current and future managers is management development systems. Control systems is the set of processes (budgeting, goal setting) and mechanisms (performance appraisal) that would encourage behavior that would help achieve organizational objectives (Flamholtz, 1995).

Developing Corporate Culture

Just as people have personalities, organizations have cultures, which are composed of shared values, beliefs, and norms. Shared values refer to the importance the organization attaches to the aspects of product quality, customer service, and treatment of employees. Beliefs are the ideas that the people in the organization hold about themselves and the firm. Lastly, the norms are the unwritten rules that guide interactions and behavior (Flamholtz, 1995).

The Model as a Whole

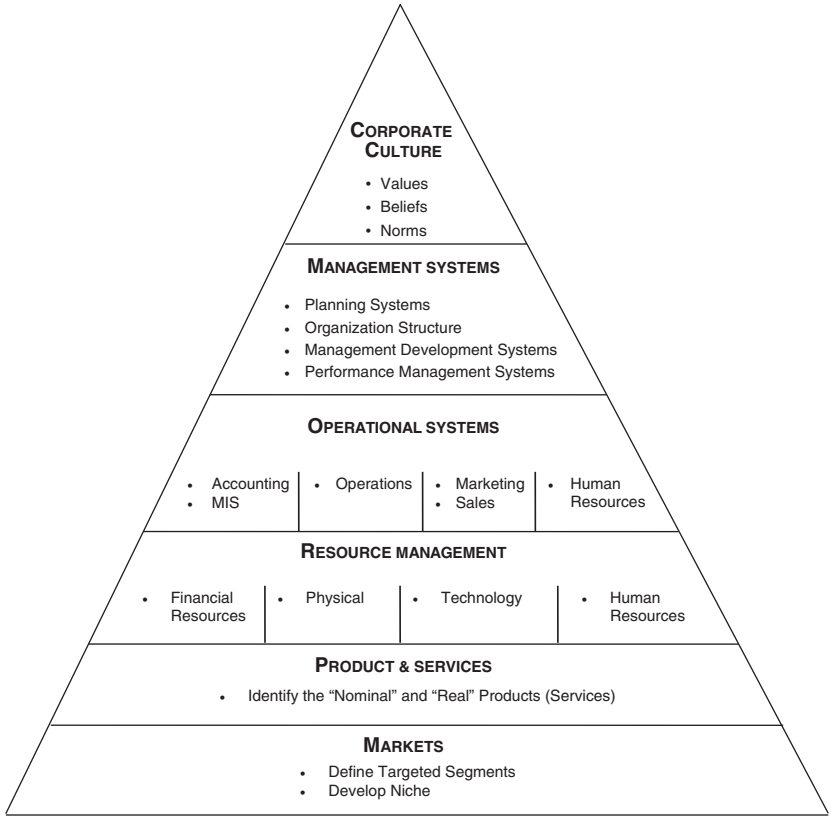
A second premise or hypotheses is that each of these tasks must be performed in a stepwise fashion in order to build a successful organization. Taken together, then, these six tasks lead to a hierarchical model of organizational development (Exhibit 1).

Similar hierarchical views are present in the previous literature. Woodward discussed a similar relation between market niche and product, and structure and culture. In addition, Chandler's (1962) book, "Strategy and Structure," suggests that a firm's structure follows from its long-term strategy.

It should be noted that the pyramid shape does not imply that the key tasks are carried out independently. All six tasks are vital for the health of the firm, and must occur simultaneously. However, the relative emphasis on each task or level of the pyramid will vary according to the organization's stage of growth (Flamholtz, 1995), as noted below.

Another hypothesis is that the top four levels of the pyramid, which form the "infrastructure" of the firm, are less susceptible to imitation (Flamholtz, 1995) and, accordingly, provide the basis for long-term sustainable competitive advantage. Thus, although competition between firms takes place at all levels, long-term sustainable advantage is primarily found at the top three levels.

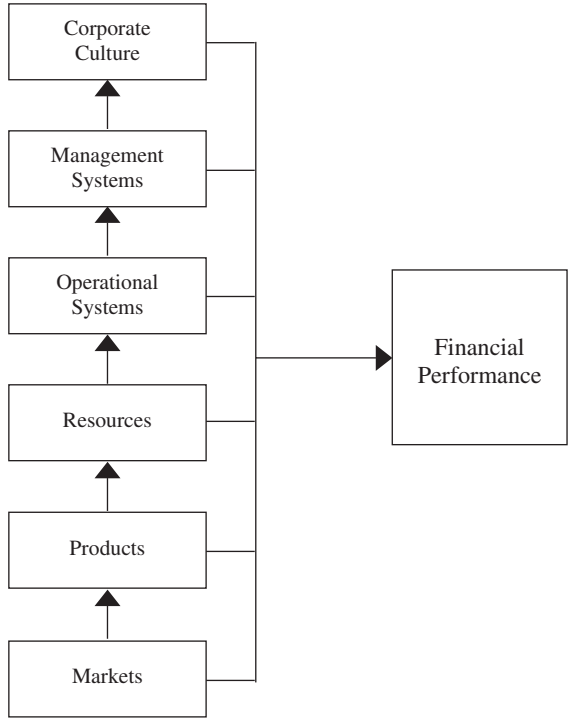
Exhibit 1. Pyramid of Organizational Development: The Six Key Building Blocks of Successful Organizations.



The Model's Variables as Drivers of Financial Performance

Another way to look at this model, and a perspective that is relevant to accounting in particular, is that the six key variables are drivers of determinants of financial performance, as shown in Exhibit 2. The set of six variables are hypothesized to account for as much as 90% of financial

Exhibit 2. Six Key Drivers of Financial Results or "Building Blocks" of Organizational Performance Success.



performance, with the remaining 10% attributable to exogenous factors. As discussed below, empirical research to date as indicated that as much as 75% of financial performance is explained by the variable in the model (Flamholtz & Kurland, 2005).

Strategic Organizational Development at Different Stages of Growth

The emphasis that should be given to each task differs depending on the size of the firm. Organizations experience developmental problems if their

Exhibit 3. Stages of Growth.

Stage	*Description*	*Critical Development Areas*	*Approximate Organizational Size (in sales: US$)*	
			Manufacturing Firms	*Service Firms*
I.	New venture	Markets and products	Less than $1 million	Less than $0.3 million
II.	Expansion	Resources and operational systems	$1 - $10 million	$0.3 - $3.3 million
III.	Professionalization	Management systems	$10 - $100 million	$3.3 - $33 million
IV.	Consolidation	Corporate culture	$100 - $500 million	$33 - $167 million

infrastructure is not consistent with their size. The parallel relationship with size and organizational structure leads to an organizational life cycle model that complements the Organizational Development Pyramid (Flamholtz, 1995), as shown in Exhibit 3.

As seen in Exhibit 3, each stage of growth is viewed as having a set of critical developmental tasks. For example, the critical tasks at Stage I are markets and products, while at Stage III the critical task is the development of management systems. Further discussion of the stages of organizational growth and the related developmental issues can be found in Flamholtz and Randle (2003).

Dysfunctional Consequences of Suboptimal Strategic Organizational Development

In the framework presented above, strategic organizational development equilibrium occurs when there is a fit between the organization's strategic development of the six key building blocks of organizational success and its size or stage of development. When this fit does not occur, the organization will experience a variety of "organizational growing pains." These growing pains are symptoms of organizational distress and an indication of the need to change, if the organization wants to continue to operate successfully.

The Classic Growing Pains

Based upon our experience in working with a wide variety of organizations, we have identified ten classic symptoms of organizational growing pains

Exhibit 4. Ten Classic Growing Pains.

TEN CLASSIC GROWING PAINS

1. People feel that "there are not enough hours in the day."
2. People spend too much time "putting out fires."
3. People are not aware of what other people are doing.
4. People lack understanding about where the firm is headed.
5. There are too few good managers.
6. People feel that "I have to do it myself if I want to get it done correctly."
7. Most people feel that "our meetings are a waste of time."
8. When plans are made, there is very little follow-up, so things just don't get done.
9. Some people feel insecure about their place in the firm.
10. The organization continues to grow in sales but not in profits.

(Flamholtz, 1995) and (Flamholtz & Randle, 2000). These growing pains, which are summarized in Exhibit 4 and described below, were derived from observations and assessments conducted with a wide variety of organizations (different sizes and industries).

(1) *People feel that* "*there are not enough hours in the day*." People feel they can work 24 h a day, 7 days a week and still not get all the required work done. When employees believe that they are being endlessly overworked, morale problems can occur. People may simply decide they can no longer operate under these conditions and may leave the organization. This will result in significant turnover costs and replacement costs related to recruiting, selecting, and training new people.
(2) *People spend too much time* "*putting out fires.*" This means that people are faced with an almost endless series of crises or "fires." Examples of "putting out fires" problems are easy to find.

"Fires" or crises were so prevalent at one $50 million manufacturing company in the U.S. that 33 managers began to refer to themselves as "fire fighters," and senior management rewarded middle management for their skills in handling crises. When it became apparent that managers who had been effective in "fire prevention" were being ignored, some of them became "arsonists" to get senior management's attention.

(3) *People are not aware of what other people are doing.* This creates a situation in which people and departments do whatever they want to do and say that the remaining tasks are "not our responsibility." Constant bickering between people over responsibility for things not getting done may ensue.

(4) *People lack understanding about where the firm is headed.* Employees may complain that "the company has no clear direction." When insufficient communication is combined with rapid changes, employees may begin to feel anxious. If anxiety increases to the point where it becomes unbearable, employees may start leaving the firm. It should be noted that turnover of this kind could be very costly to the company.

(5) *There are too few good managers.* Although the organization may have many people who hold the title of "manager," it may not have good or effective managers. Rapid growth at Apple Computer led Steven Jobs to bring in "professional managers" to help manage the company because it had not developed a cadre of managers as it grew. However, this led to the inevitable culture clash, and to Jobs' resignation.

(6) *People feel that "I have to do it myself to get it done correctly."* Increasingly, as people become frustrated by the difficulty of getting things done in an organization, they come to feel that "if I want to get something done correctly, I have to do it myself." Operating under this mindset departments become isolated from one another and teamwork becomes minimal.

(7) *Most people feel "our meetings are a waste of time."* Unfortunately, at many companies, meetings have typically no planned agendas, and often they have no designated leader. As a consequence, the meetings become a free-for-all, tend to drag on interminably, and seldom result in decisions.

Other complaints about meetings involve lack of follow-up on decisions that are made. Meetings are also ineffective if people ignore the goals that have been set or fail to monitor their progress toward these goals.

(8) *When plans are made, there is very little follow-up so things just don't get done.* Recognizing that the need for planning is greater than in the past,

a CEO may introduce a planning process. People go through the motions of preparing business plans, but the things that were planned just do not get done. In some cases, there is no follow-up because the company has not yet developed systems adequate to monitor its goals. In other cases, follow-up does not occur because personnel have not received proper training in setting, monitoring, and evaluating goals.

(9) *Some people feel insecure about their place in the organization.* Sometimes the board has become anxious about problems facing the organization and has therefore hired a "heavy-weight" manager from outside. This action may have been accompanied by the termination of one or more current managers. Employees begin to wonder whether they will be the next to "get the axe." In an attempt to protect themselves, they keep their activities secret and do not "make waves." This results in isolation and a decrease in teamwork. When anxiety becomes too high, it may result in morale problems, turnover, or a very political environment.

(10) *The organization continues to grow in sales but not in profits.* If all the other growing pains are permitted to exist, this final symptom may emerge. In some instances, sales continue to increase while profits remain flat, so that the company is succeeding in only increasing its workload. In the worst cases, sales increase while overall profits decline.

This set of classic growing pains are not only problems in and of themselves; they are symptoms of a deeper problem, and a "signal" or warning that the organization needs to make a fundamental change in its infrastructure, as explained below. Although it is tempting to look at growing pains from a binary ("yes" or "no") perspective, as we shall explain below, it is more useful to view them on a continuum, i.e., the degree to which they exist in a particular organization.

Nature and Causes of Organizational Growing Pains

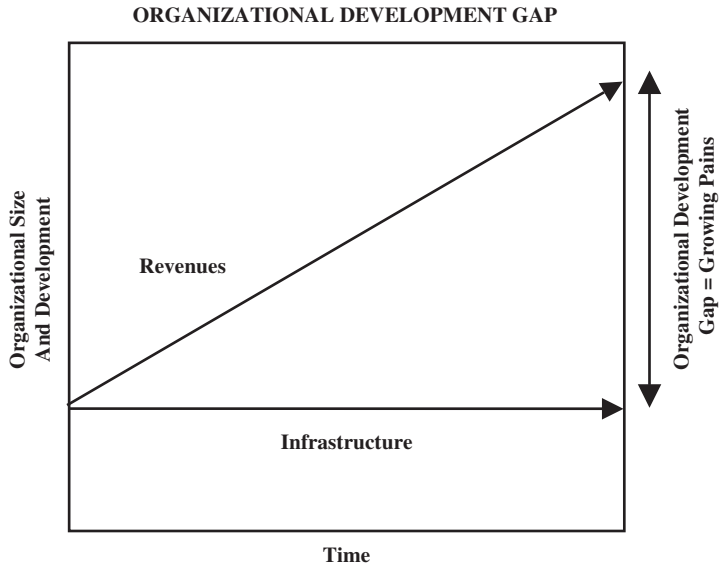
Growth, though essential to organizations over the long term, creates its own set of problems: the growing pains described above. These growing pains are symptoms that something has gone wrong in the growth and development of a business enterprise. They are a symptom of organizational

distress, and an early warning or leading indicator of future organizational difficulties, including financial difficulties.

Growing pains indicate that the "infrastructure" of an enterprise (i.e., the internal operational and management systems it needs at a given stage of growth) has not kept up with its size, as measured by its revenues. For example, a business with $200 million (U.S.) in revenues may only have an infrastructure to support the operations of a firm with $50 million in revenues, or one-fourth its size. This type of situation typically occurs after a period of growth, sometimes quite rapid growth, where the infrastructure has not been changed to adjust to the new size and complexity of the organization. The result, as shown in Exhibit 5, is an "organizational development gap" (that is, a gap between the organization's actual infrastructure and that required at its current size or stage of development), which produces the growing pains.

As a rule of thumb, whenever an organization doubles in size (as measured by its revenues), it is essentially a different company and requires a new

Exhibit 5. Organizational Development Gap.



infrastructure to support its operations. If the infrastructure has not been adjusted to reflect the increased size, a variety of classic growing pains will be experienced.

Growing pains can and do occur in organizations of all sizes, including the largest industrial enterprises. However, they are mostly characteristic of early stage entrepreneurial companies, even in those organizations where revenues exceed $1 billion. Although growing pains are the result of organizational success (i.e., successful development of a market and product), they can lead to great difficulties and even foreshadow failure. For example, Osborne Computers, a pioneer in the portable "personal" (micro) computer business, achieved $100 million in revenues after being in business for only 2 years, but went into bankruptcy in the 3rd year.

SUMMARY OF THE FRAMEWORK'S IMPLICATIONS

Several implications can be derived from the framework described above. These are summarized below:

(1) The initial premise or implication from this framework is that organizations must perform certain tasks to be successful at each stage of their growth.
(2) A second premise is that each of these tasks must be performed in a stepwise fashion in order to build a successful organization.
(3) Another implication is that the top four levels of the pyramid, which form the "infrastructure" of the firm, are less susceptible to imitation (Flamholtz, 1995), and, accordingly, provide the basis for long-term sustainable competitive advantage.
(4) Each stage of growth is viewed as having a set of critical developmental tasks. For example, the critical tasks at Stage I are markets and products, while at Stage III the critical task is the development of management systems.
(5) Strategic organizational development equilibrium occurs when there is a fit between the organization's strategic development of the six key building blocks of organizational success and its size or stage of development. When this fit does not occur, the organization will experience a variety of "organizational growing pains." These growing pains are symptoms of organizational distress and an indication of the need to change, if the organization wants to continue to operate successfully.

EMPIRICAL RESEARCH TO ASSESS VALIDITY OF FRAMEWORK

To assess the validity of the framework presented above and to provide empirical support for its proposed implications, Flamholtz and colleagues have been engaged in a program of empirical research. In the following section, the empirical research to date to test the model's predictive validity and its related hypotheses will be summarized.

Strategic Organizational Development and Financial Performance

Flamholtz and Aksehirli (2000) proposed a link between the organizational success model and the financial success of organizations. To test this hypothesized relationship, they analyzed financial and non-financial information relevant to the hypothesized model for eight pairs of companies in different industries. Each company was evaluated in terms of the six key strategic building blocks, and scores were assigned to indicate the degree of the organization's development. Average Return on Equity was used as an indicator of financial performance. Using the Friedman Two-way Analysis of Variance and a regression analysis, they found a statistically significant relationship between the proposed model of organizational success and financial performance.

The major implication of this research is that it provides empirical support for the use of the Pyramid Model as a managerial tool, as we proposed in the previous section.

Strategic Organizational Development and Financial Performance: Additional Evidence

In addition, Flamholtz and Hua (2002a) report the results of an empirical test of the hypothesized relationship regarding financial success and the degree of development of six key variables (or "strategic building blocks") included in the organizational development pyramid within a *single* firm. The research site was a U.S.-based, medium-sized industrial enterprise. The company is a parts manufacturer for industrial truck, and other automotive businesses. It is a supplier of parts for such companies as Ford Motor Company, Navistar, and Dana Corporation.

To assess this issue, they compared divisional data, the degree of organizational development, with divisional "EBIT" (earnings before interest and taxes), a classic measure of financial performance for 18 divisions. Specifically, they ran a regression between: (1) the degree to which each division was perceived as being developed on the six key strategic building blocks as a whole (i.e., the average pyramid development score), and (2) EBIT. This regression was statistically significant. This result supports the hypothesis of a relationship between the degree of strategic organizational development and the financial performance of organizations.

Another question concerned the thresholds of strategic organizational development for profitability of individual companies or operating units. Specifically, they wanted to identify potential "benchmarks" of organizational development to serve as guideposts for developing the six key strategic building blocks. Stated differently: What are the levels of strategic organizational development required for profitability and superior profitability in companies?

They found that all of the six divisions with strategic organizational development scores greater than 3.0 were profitable. In contrast, for the nine divisions with strategic organizational development scores less than 3.0, six were profitable and three were "unprofitable" (i.e., negative EBIT).

This study has implications for the level of strategic organizational development required for optimal profitability. One major of this study is that it provides additional empirical support for the use of the pyramid modes as proposed earlier in this article. Another major managerial implication of this study is that there is a high (in this study 100%) probability of profitability for organizations with Pyramid scores greater than 3.0. Similarly, it also suggests that there is a 33% chance of being unprofitable for organizations with Pyramid scores less than 3.0. While a level of development of 3.0 seems to be the threshold for being profitable, most organizations want to achieve superior financial performance.

Strategic Organizational Development and Financial Performance: Additional Evidence

Flamholtz and Kurland have replicated the study reported above in another, different kind of firm, a diversified financial institution (Flamholtz & Kurland, 2005). The paper also provides a test of the notion that an organization's infrastructure (defined below) contains the key drivers of an organization's performance and profitability. Specifically, it examines the

hypothesis that the key determinants of organizational profitability are the resources, operational systems, management system, and culture of an organization (infrastructure). If valid, this has implications for competitive strategy as well as strategic organizational development Flamholtz and Hua with the assistance of Aksehirli (2003).

The research site was a company called ''Diversified Financial Corporation'' (a disguised name). Diversified Financial Corporation is a large, U.S.-based, financial institution. It consists of several (15) subsidiary companies, with one operating in the U.K. In 2001, company had more than $3.5 billion in revenues.

The 15 divisions are comprised of two groups: (1) mortgage-related businesses and (2) other financial businesses. There are seven financial business units in the second group, and these comprised the sites for the present study. The similarities between the divisions present a relatively unique opportunity for comparison. Each of the seven individual companies, or ''divisions,'' as they were termed, operated in various parts of the United States.

The methodology used was the same as in Flamholtz and Hua (2002a). Specifically, the senior executives of each division of the company were asked to rate each division on each of the six key strategic building blocks (markets, products, etc.) using a five-point Likert scale. The results of this assessment were used to construct an ''average pyramid development score,'' a measure of the ''average strategic organizational development'' of each division.

The possible scores range from 1.0 to 5.0, where 1.0 is the lowest possible score and 5.0 is the highest possible score. To measure financial performance or the ''bottom line'' for each division we used a measure of divisional performance (i.e., divisional ''profit margin,'' a measure of gross margin) that was reported throughout the company on a quarterly basis, one that Diversified Financial Corporation uses to assess divisional performance, for these divisions.

The ''average divisional strategic organizational development'' (a measure of the degree of strategic development of each of the divisions in terms of the six key strategic building blocks of successful organizations, as measured by a five-point Likert scale), was used as an input into a regression equation. The regression equation describing the relationship among variables is: y (Profit) $= -311.18 + 107.65 *$ Pyramid Score. Adjusted R^2 is 0.73, and is statistically significant at 0.02 level ($F = 14.66$). This means that approximately 73% of PROFIT are explained by the six key factors. This result provides strong support for the hypothesis of a relationship between the

degree of strategic organizational development and the financial performance of organizations.

CORPORATE CULTURE AND FINANCIAL PERFORMANCE

In addition to the overall tests of the strategic organizational development model, there has also been an empirical test of the effects of corporate culture on financial performance (Flamholtz, 2001). "Corporate culture" is one of the six key building blocks included in the pyramid framework. It is also hypothesized to be the critical developmental factor at Stage IV (see Exhibit 2).

Previous authors (Kotter & Hesket, 1992) have suggested that culture has an impact on financial performance. Unlike previous studies, which have only examined the effects of culture on financial performance using cross-sectional data, Flamholtz (2000) did a study of the impact culture has on financial performance in a single organization.

The study involved developing statements describing the core values of the desired culture of the company as a whole, as well as determining the extent to which the divisions' culture was consistent with the stated desired culture. This was measured by using a survey with a Likert scale (Flamholtz, 2000). These data were then used as an input to address the question concerning the impact of corporate culture on financial performance. The hypothesis was that the greater the degree of agreement of the divisional culture with the overall desired corporate culture, the greater financial performance. Financial performance was measured as EBIT.

The results, using a regression analysis, indicate that there is a statistically significant relationship between culture and financial performance (measured by "EBIT," or earnings before interest and taxes). Thus, these results provide support for the previously hypothesized relationship between culture and financial performance, with significant implications for management theory and practice.

One of the major implications concerns the potential sources of competitive advantage. One of the hypotheses is that the top four levels of the pyramid, which form the "infrastructure" of the firm, are less susceptible to imitation (Flamholtz, 1995), and, accordingly, provide the basis for long-term sustainable competitive advantage. Culture is one of the key components of organizational infrastructure, and if there are demonstrable

differences in culture across business units, which are associated with differences in profitability, this provides support for the notion that organizations compete not only in products and markets but in infrastructure as well.

In another study, Flamholtz and Kannan-Narasimhan (2005) studied the differential impact of key elements or components of culture upon financial performance. Using multiple regressions and factor analysis, they identified six statistically significant cultural factors.

Infrastructure and Competitive Advantage

One of the hypotheses presented above is that the top four levels of the pyramid, which form the "infrastructure" of the firm, are less susceptible to imitation (Flamholtz, 1995), and, accordingly, provide the basis for long-term sustainable competitive advantage. At present there is no published research on this issue. However, Flamholtz et al. (2003) have conducted research on this issue. They have found empirical support for this hypothesis.

The major implication of this study is that it challenges the convention paradigm of strategy, which focuses almost exclusively upon external forces. The research by Flamholtz et al. (2003) indicates that competitive advantages can occur within "the black box" systems internal to an organization.

Growing Pains and Financial Performance

As discussed above, when an organization grows, it will almost inevitably experience a classic set of "growing pains." These growing pains are "symptoms" that something has gone wrong in the process of strategic organizational development, and an "early warning" of significant future problems. More specifically, strategic organizational development equilibrium occurs when there is a fit between the development of the six key building blocks of organizational success and the organization's size or stage of development (Flamholtz, 1995). When this fit does not occur, the organization will experience a variety of "organizational growing pains."

Flamholtz and Hua (2002b) performed an empirical test of the hypothesized relationship between "organizational growing pains" and corporate financial performance. They also addressed the question: are there benchmark levels of growing pains which might be used to predict which organizations will be profitable versus those which are likely to be unprofitable?

Previous to this research, the hypothesized relationship between growing pains and performance in previous literature has been conceptual in nature; in contrast, this study presents some very specific "benchmarks" for growing pains in relation to successful organizational financial performance.

To study whether there is a statistically valid predictive relationship between growing pains and EBIT, they calculated a regression equation based upon these two variables. The results of this statistical test indicate that there is a statistically significant relationship between growing pains as a predictor of EBIT. This means that growing pains are a predictor of financial performance or the "bottom line" (EBIT).

An analysis of the relationship between specific growing pains scores and financial performance was also conducted to determine benchmark levels of "safe" versus "unsafe" growing pains. The results suggest that there appears to be a maximum level of growing pains beyond which organizational financial health is at risk. This suggests that there is a "maximum healthy growing pains score" to provide the highest probability of success, and confirms that there do appear to be thresholds levels of growing pains, which might be used to predict which organizations will be profitable versus those which are likely to be unprofitable.

The data derived from this study provide empirical support for the notion that growing pains have an impact on financial performance, and that there are threshold levels of growing pains that are "unsafe" or "unhealthy" for future financial performance. The results of the analysis suggest that there is a (very strong) statistically significant relationship between growing pains and financial performance.

The major implication of this research is that there appears to be a maximum level of growing pains beyond which organizational financial health is at risk. Specifically, the maximum "healthy" level of growing pains appears to be "32." This means that to optimize the chances of being profitable an organization ought to keep its growing pains score less than 32.

IMPLICATIONS FOR MANAGEMENT, BOARDS, AUDITORS, AND RESEARCHERS

The proposed model and empirical research findings presented above have significant implications for accounting, information, and control as well as for management theory and practice. The specific implications of each individual research study have been stated above. However, this section

examines the broader implications of the model and research as a whole for management, boards, auditors, accountants, and researchers.

Strategic Organizational Development and Financial Performance

The data derived from the set of empirical studies surveyed above provide an indication that the proposed model of strategic organizational development does have an impact on financial performance. This has important implications for management theory and practice. It is one thing to assert that organizational development is a significant factor of organizational success and quite another to be able to demonstrate that the effective management of these variables can enhance profitability.

Managers can have confidence in using the framework to assess the strategic development of their companies as well as to plan for its future development. This suggests that the strategic planning process ought to be based upon the pyramid as a ''strategic lens'' for the development of organizations. Although all six functions that make up the pyramid should be managed successfully in order to achieve good financial performance, practitioners can incorporate the organizational life-cycle model to decide which tasks to emphasize at each stage of growth.

Another implication for management is the fact that the organizations are competing at each level of the pyramid. Since markets can be easily entered and products can be easily copied, the real competition goes on at the top four levels of the pyramid. This is the area where organizations can develop sustainable competitive advantages.

Corporate Culture and Financial Performance

The data derived from the set of empirical studies surveyed above provides an indication that the corporate culture does have an impact on financial performance. This has important implications for management theory and practice. It is one thing to assert that corporate culture is a significant factor of organizational success and quite another to be able to demonstrate that the effective management of this variable can enhance profitability.

One of the major implications concerns the potential sources of competitive advantage. One of the hypotheses is that the top four levels of the pyramid, which form the ''infrastructure'' of the firm, are less susceptible to imitation (Flamholtz, 1995), and, accordingly, provide the basis for

long-term sustainable competitive advantage. The data indicate that there is a statistically significant relationship between culture and financial performance (measured by "EBIT," or earnings before interest and taxes). Thus, these results provide support for the previously hypothesized relationship between culture and financial performance.

Culture is one of the key components of organizational infrastructure, and since there are demonstrable differences in culture across business units, which are associated with differences in profitability, this provides support for the notion that organizations compete not only in products and markets but in infrastructure as well. Culture, then, is a potential source of competitive advantage, and, in turn, differential financial performance.

Growing Pains and Financial Performance

In addition, as we have seen, organizational growing pains can directly influence financial performance or the so-called "bottom line." As a result, management needs (1) to understand the nature and causes of growing pains, (2) to have a method of measuring them, (3) a template to assess their severity, and (4) a strategy for managing them.

Variations exist, but it is clear that organizations of all sizes and types experience some growing pains. Severity of these problems can be affected by the rate of growth experienced by the organization. Managers of rapidly growing companies of any size or type must learn to recognize organizational growing pains and take steps to alleviate them so that their organizations can continue to operate successfully. The payoff will reduce growing pains and an increased likelihood of a positive "bottom line."

What should an organization do to minimize or avoid the problems associated with growing pains? Most entrepreneurs are always concerned with the risk of failure if revenues are insufficient to cover expenses. However, many ignore the equally damaging risks of choking on their own rapid growth. To avoid the problems accompanying hyper-growth, a company must have an infrastructure that will absorb that growth. If a company anticipates rapid growth, then management must invest in building the required infrastructure before it is actually necessary. It is very difficult, and sometimes impossible, to "play catch-up" with organizational infrastructure. Some companies, such as Starbucks Coffee, Compaq Computer, and PacifiCare had a strategy of having their infrastructure in place prior to their explosive growth and reaped the benefits of this investment. In contrast, Boston Markets, Osborne Computers, and MaxiCare, did not have their

infrastructure in place prior to explosive growth and all three have experienced bankruptcy. Thus, the ideal strategy for a firm that anticipates rapid growth is to build an infrastructure sufficient for the size of the organization it anticipates becoming, prior to actually reaching that size.

This strategy of building the infrastructure prior to growth is not merely appropriate for large companies, but for relatively small entrepreneurships as well. For example, several years ago, one of the authors met with the president of a U.S. service firm specializing in insurance-based benefit programs for executives when the firm had approximately $3 million in annual revenues. At that time, the authors of this article advised the CEO that it was probably premature to build the infrastructure to the extent that was being contemplated. However, the CEO indicated that he wanted his firm to grow to $50 million in revenue within five years. He then proceeded to invest in building the infrastructure of his company before it was actually necessary. This was a wise move, because the company actually grew to more than $65 million in revenue within 5 years.

Given the research findings about growing pains cited above, it appears that growing pains can be used as leading indicators of future financial performance. The U.S. Federal Reserve monitors leading indicators of economic activity to predict the direction of GNP and inflation. Similarly, growing pains might be used as leading indicators of future changes in organizational financial performance. In addition, our findings concerning the maximum level of growing pains in relation to the levels of profitability are, at a minimum, suggestive of the need to control or at least minimize growing pains.

Implications for Boards of Directors and External Auditors

These findings also have implications for Boards of Directors and external auditors. Recent experiences in the U.S., with Enron, Waste Management, and other publicly traded enterprises suggest the need for improved methods of control (Flamholtz, 1996; Nilsson & Olve, 2001). There are complex issues involving the balance of power among management, boards, and auditors not only in the U.S., but throughout Europe and Asia as well (Hooghiemstra & Van Manen, 2002). What are required are tools that can help identify potential problems before they occur.

Since growing pains can be measured and we have shown that they are clearly linked to financial performance, it would be useful to report growing pains to the board. This would be done on a comparative basis across time.

Independent auditors might also find this information useful as a signal to look for organizational problems.

Implications for External Financial Reporting

It is increasingly recognized that conventional accounting does not do an adequate job of reflecting the economic realities of a business. The current accounting model was developed for the agricultural and industrial era, and, accordingly, it does not do an adequate job of reflecting the impact of intangibles and intellectual property on financial statements. As Flamholtz and Main (1999) have indicated, the market value of companies such as Microsoft, Intel, and Amgen greatly exceeds their book value. The market is telling us that there is a great value in unmeasured and unreported intangible assets.

The research data reported in this paper have supported the market's conclusions. It has indicated that an organization's infrastructure (consisting to a great extent of intangibles and intellectual property, such as operational systems, management systems, and culture) can account for about 73% of profit (Flamholtz & Kurland, 2005). Culture alone has been shown to account for about 46% of EBIT (Flamholtz, 1998).

This suggests that accountants need to begin to report this type of information both as a supplement to conventional financial reports and in internal management accounting systems. This, in turn, implies a broader role for accounting but it is not inconsistent with other calls for a redefinition of accounting for many decades.

Implications for Internal Management Accountants

These findings also have implications for internal management accountants. They speak to the kinds of information, which management accountants ought to be providing to management. Previous research has suggested that accounting control systems need to be viewed in a larger organizational context (Flamholtz, 1983, 1996). If the data concerning strategic organizational development explain or account for as much as 55–73%, or even more of financial performance, this data ought to be monitored and reported on an ongoing basis.

In addition, these results challenge the perspectives used in the so-called "Balanced Scorecard" developed by (Kaplan & Norton, 1992, 1993,

1996a,b). Instead of the four "perspectives" proposed by Kaplan and Norton, *for which no empirical evidence has ever been presented*, the proposed model can be used with greater confidence as a valid set of indicators of the health of an enterprise as part of a revised Balanced Scorecard (Flamholtz, 2003).

FUTURE RESEARCH

From an academic perspective, the results reported here are preliminary but promising. The results of the research surveyed here represent the first attempt in the empirical analysis of organizational development pyramid framework and should be supplemented with further studies. It would be valuable for future research to replicate the current study, not only in North American environment but in Europe and Asia as well.

This paper also suggests that the level of strategic organizational development, as well as the level of growing pains, can be used to estimate the future financial success of the firm. Although the results reported here are promising, it remains for future research to examine this phenomenon with a longitudinal study using time series analysis.

In additional as yet unpublished research, we have used a combined measure of strategic organizational development and growing pains as a measure of "organizational health" or "risk." We have found that there is a very strong statistically significant relationship between this measurement of organizational health/risk and financial performance. Specifically, Adjusted R^2 is 0.63, and is statistically significant at 0.00045 level. This means that approximately 63% of EBIT is explained by the combined measure (strategic organizational development and growing pains). This result provides strong support for the hypothesis of a relationship between the degree of strategic organizational development and growing pains and the financial performance of organizations.

Within accounting and information systems per se, empirical studies of the application and use of the proposed model as components of a control system need to be investigated. At present, the author has applied this model as part of a comprehensive strategic management system at many companies, including Starbucks Coffee Company, Countrywide Financial Corporation, PacificCare, American Century Investors, Navistar, and many other organizations (Flamholtz & Randle, 1998, 2000). However, these can be supplemented by additional applications, especially in the European and Asian context.

CONCLUSION

The organizational development pyramid framework can be a promising tool in predicting the future financial performance of the companies. In combination with stages of growth, the organizational development pyramid can be used to assess a company's success in fulfilling the critical tasks for each stage of growth. In addition, as we have seen, organizational growing pains can directly influence financial performance or the so-called "bottom line."

This has significant implications for accounting, organizations, and society. It speaks to the need to supplement the information provided by accountants to management as well as to change the components of organizational control systems. This is consistent with previous research (Flamholtz, 1983). It also speaks to the need for this information by Boards of Directors.

This framework offers the basis of a different paradigm of organizational success and failure for organizations at different stages of growth, from new entrepreneurships to established companies. Although the research is not definitive, it offers some promising findings and opens the way to new questions.

ACKNOWLEDGMENT

The author gratefully acknowledges that the research reported in this paper was partly supported by the Harold Price Center for Entrepreneurial Studies in the Anderson School at UCLA.

REFERENCES

Aldrich, I. (1979). Organizational passages: Diagnosing and treating life cycle problems in organizations. *Organizational Dynamics*, *8*(1), 3–24.

Brittain, J. W., & Freeman, J. (1980). Organizational proliferation and density-dependent selection. In: J. R. Kimberly & R. H. Miles and associates (Eds), *The organizational life cycle issues in the creation, transformation, and decline of organizations* (pp. 291–338). San Francisco: Jossey-Bass.

Burns, T., & Stalker, G. M. (1961). *The management of innovation*. London: Tavistock.

Carroll, G. R., & Yangchung, P. H. (1986). Organizational task and institutional environments in ecological perspective: Findings from the local newspaper industry. *American Journal of Sociology*, *91*, 838–873.

Chandler, A. D. (1962). *Strategy and structure: Chapters in the history of the industrial enterprise*. Cambridge: MIT Press.

Child, J., & Keiser, A. (1981). Development of organizations over time. In: P. C. Nystrom & W. H. Starbuck (Eds), *Handbook of organizations: Adapting organizations to their environments* (pp. 28–64). New York: Oxford University Press.

Flamholtz, E. (1983). Accounting, budgeting, and control system in their organizational context: Theoretical and empirical perspectives. *Accounting, Organizations, and Society*, *8*(2/3), 153–169.

Flamholtz, E. (1995). Managing organizational transitions: Implications for corporate and human resource management. *European Management Journal*, *13*(1), 39–51.

Flamholtz, E. (1996). *Effective management control: Theory and practice*. Norwell, MA: Kluwer Academic Press.

Flamholtz, E. (2001). Corporate culture and the bottom line. *European Management Journal*, *19*(3), 268–275.

Flamholtz, E. (2003). Putting validity and balance into the balanced scorecard. *Journal of Human Resource Costing & Accounting*, *7*(3), 15–26.

Flamholtz, E. G., & Aksehirli, Z. (2000). Organizational success and failure: An empirical test of a holistic model. *European Management Journal*, *18*(5), 488–498.

Flamholtz, E., & Hua, W. (2002a). Strategic organizational development and the bottom line: Further empirical evidence. *European Management Journal*, *20*(1), 72–81.

Flamholtz, E., & Hua, W. (2002b). Strategic organizational development, growing pains and corporate financial performance: An empirical test. *European Management Journal*, *20*(5), 527–536.

Flamholtz, E., & Hua, W. with the assistance of Aksehirli, Z. (2003). Searching for competitive advantage in the black box. *European Management Journal*, *21*(2), 222–236.

Flamholtz, E., & Kannan-Narasimhan, R. (2005). Differential impact of cultural elements on financial performance: An empirical analysis. *European Management Journal*, *23*(1), 50–64.

Flamholtz, E., & Kurland, S. (2005). *Strategic organizational development, infrastructure, and financial performance: An empirical investigation*. Unpublished Working Paper to be presented at the UCLA/ Oxford University conference on research in entrepreneurship, (June), 22–24.

Flamholtz, E., & Main, E. (1999). Current issues, recent advancements, and future directions in human resource accounting. *Journal of Human Resource Costing & Accounting*, *4*(1), 11–20.

Flamholtz, E., & Randle, Y. (1998). *Changing the game: Organizational transformations of the first, second, and third kinds*. New York: Oxford University Press.

Flamholtz, E., & Randle, Y. (2000). Growing pains: Transitioning from an entrepreneurship to a professionally managed firm (New, rev. ed.). San Francisco: Jossey-Bass.

Flamholtz, E., & Randle, Y. (2003). Measuring and managing organizational scale-up. Presented at "*Conceptualizing Scale-up: Multidisciplinary Perspectives*," University of Chicago Data Research and Development Center, Washington, DC, November 3–4. To be published in B. Schneider, L. Hedges & S. K. McDonald (Eds), *Conceptualizing scale-up: Multidisciplinary perspectives*, (forthcoming), 2005.

Freeman, J., & Hannan, M. T. (1983). Niche width and the dynamics of organizational populations. *American Journal of Sociology*, *88*, 1116–1145.

Hooghiemstra, R., & Van Manen, J. (2002). Supervisory, directors, and ethical dilemma: Exit or voice. *European Management Journal*, *20*(1), 1–9.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard – measures that drive performance. *Harvard Business Review*, *70*(1), 71–79.

Kaplan, R. S., & Norton, D. P. (1993). Putting the balanced scorecard to work. *Harvard Business Review*, *71*(5), 134–147.

Kaplan, R. S., & Norton, D. P. (1996a). *The balanced scorecard*. Boston, MA: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. (1996b). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, *74*(1), 75–85.

Kotter, J. P., & Hesket, J. L. (1992). *Corporate culture and performance*. New York: The Free Press.

Midgley, D. F. (1981). Toward a theory of the product life cycle: Explaining diversity. *Journal of Marketing*, *45*, 109–115.

Nilsson, F., & Olve, N. (2001). Control systems in multi-national companies: From performance management to strategic management. *European Management Journal*, *19*(4), 344–358.

Peters, T. J., & Waterman, R. H. (1982). *In search of excellence*. New York: Harper & Row.

Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: A resource dependence perspective*. New York: Harper & Row.

Starbuck, W. (1965). Organizational growth and development. In: J. G. March (Ed.), *Handbook of organizations* (pp. 451–533). Chicago: Rand McNally.

Tushman, M. L., Virany, B., & Romanelli, E. (1985). Executive succession, strategic reorientation, and organization evolution: The mini computer industry as a case in point. *Technology in Society*, *7*, 297–313.

Walton, R. E. (1986). A vision-led approach to management restructuring. *Organizational Dynamics*, *14*(4), 9–16.

This page intentionally left blank

THE PYRAMID OF ORGANIZATIONAL DEVELOPMENT AS A PERFORMANCE MEASUREMENT MODEL

K. J. Euske and Mary A. Malina

In recent times, performance measurement has moved from lists of key performance indicators to more comprehensive business models that reflect the firm as a system. Consistent with this more comprehensive approach, Flamholtz (2005) presents a holistic performance measurement model termed the Pyramid of Organizational Development. The Pyramid presents six key building blocks of successful organizations: (1) markets, (2) products, (3) resources, (4) operational systems, (5) management systems, and (6) corporate culture. Flamholtz suggests that different levels of the Pyramid are relatively more important at different stages of company growth. He argues that if fit between the Pyramid and growth stage is not achieved, then the organization will experience growing pains that negatively impact financial performance. Our task is to comment on how to improve and build upon this model, as presented in Flamholtz (2005), with an eye to the more general question of what we should expect of performance measurement

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 167–175

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14007-6

models.[1] We proceed with a discussion of model characteristics, followed by model testing, and then implications for such models.

MODEL CHARACTERISTICS

Capturing Complex Interrelationships

If a comprehensive performance model for business is to be useful in an analytic and predictive sense, the model must capture the interrelationships of factors that influence organizational performance such as organizational maturity, size, products and services, management systems, industry characteristics, and environmental influences. Flamholtz includes a number of key factors in his model. However, Flamholtz's explication of the factors does call into question some aspects of the model. For instance, Flamholtz explicitly equates level of sales revenue with specific growth stages of the organization and implicitly equates level of sales revenue with the maturity of the organization. Although these factors may be correlated in many organizations, care must be taken so that the comprehensive performance model does not confound key factors.

A manufacturing organization, such as a shipyard, could be a relatively new venture with only one order and be in the highest sales revenue category of the Flamholtz model. On the other hand, a firm could be very mature with a small sales volume. It appears that the model as presented by Flamholtz (2005) is meant specifically to apply to organizations where dollar sales volume categories as shown in Exhibit 3 (Flamholtz, 2005) correlate with both the chronological age of the organization and the maturity of its products and processes. Greiner (1998) clearly distinguishes between size and age in his model of organizational evolution. Interestingly, his definition of size is vague but he does discuss both number of employees and sales volume as indicators of growth. Growth and organizational maturity are complex concepts that a rich holistic performance measurement model needs to fully capture. In Flamholtz and Randle (2005), the authors do discuss the complexity of the relationships and the difficulty of operationalizing these concepts.

Assessing Issues of Use Versus Design

The applicability of a performance measurement model will also depend on its ability to identify and relate issues to the design versus the use of the

performance factors. It is possible to have proper processes and systems in place and not use them well, just as it is possible to have well utilized but weak processes and systems. In the Flamholtz model, a consequence of misfit between the Pyramid and firm growth stage is labeled growing pains. It is unclear whether fit, or lack thereof, is based on the design of the building blocks, how employees use the building blocks, or both. The logic behind the model appears to be that it is designed for relatively large growing organizations that render the infrastructure of the organization framework obsolete at regular intervals. An implicit assumption in the model seems to be that the issues are those of design not in use. A generalizable holistic performance model would address not only the growing organization, but also those that achieve a steady state before reaching the categories containing the larger-sized organizations listed in Exhibit 3 (Flamholtz, 2005). Such a model would more readily support the analysis of issues of both design and use.

Reconciling Divergent Views

A generalizable holistic performance measurement model will need to address the seemingly divergent views regarding factors that are most likely to be important to an organization at various stages of its growth. For instance, several parallels can be made between the Pyramid and Simons' (1995) Levers of Control model. As an example, Simons' beliefs and boundary systems mirror the corporate culture level of the Pyramid, while Simons' interactive and diagnostic control systems are similar to the Pyramid's management systems level. Simons addresses the concept of fit between his levers of control and life cycle stages. Simons suggests that beliefs and boundary systems should be implemented as a firm begins to expand. New locations, new product offerings, and an increase in the number of employees necessitate top management formally document and communicate the values, beliefs, and norms of the organization. However, Flamholtz (2005) suggests that this is optimally performed later at the consolidation phase. The implications for the analysis of poor or even well-performing systems differ depending on the model adopted.

Specifying Causal Relationships

Few would argue with the observation that past experience conditions our reactions to the future. We are likely to use or adapt past successful

intervention strategies to address new experiences. Eventually, as Greiner (1998) argues, the very practices that were successful in the smaller and younger organization become a problem as the organization grows and matures. If we could successfully judge when the old practices and structures are a problem, we could replace them with practices and structures that appear to be appropriate. However, if we misjudge what needs to be replaced, the fix could in effect become the problem.

The issue of inappropriate adjustments must be incorporated in any holistic performance measurement model. Otherwise, the direction of causality in the performance measurement model along with its usefulness as a tool to enhance performance will be open to question. For instance, in Flamholtz's model, the assumption that the misfit between sales revenue and organizational infrastructure causes growing pains is tenuous. In order to mitigate the growing pains, Flamholtz suggests that organizations put their larger, improved infrastructures in place prior to anticipated growth. However, putting the larger, improved infrastructure in place may result in the growing pain described as growth in sales but not in profits. This growing pain could result from changing the infrastructure too soon, not too late. Perhaps, so much money was invested in improving infrastructure that current profits suffered. There may be circularity in the causal cycle of growing pains and infrastructure.

Defining the Degree of Generalizability

It is an open question whether performance measurement models are unique to each organization or are generalizable across companies. To help ensure appropriate application, performance measurement models should be defined in terms of their generalizability. For example, a balanced scorecard (Kaplan & Norton, 1996, 2001) is virtually unique to each organization since it is tailored to each company's specific strategy. Cross-sectional assessment of performance using balanced scorecards is nearly impossible. The Pyramid, however, has potential to assess performance across companies. Scales used to assess the level of organizational development in early growth stages appear to be rather generic. For example, Flamholtz and Hua (2002a) assess the level of organizational development based solely on six questions, one for each building block. However, once a company moves past the early growth stages, Flamholtz suggests that competitive advantage becomes rooted in the company's unique culture. At that point, cross-sectional assessment of performance becomes less plausible. Idiosyncrasies of

firm-level definitions of organizational culture are likely to emerge making cross-sectional comparisons difficult. For example, Flamholtz (2001) assesses the level of cultural development, just one of the six building blocks, based on more than 25 questions developed specifically to map to that particular organization's strategy. The Pyramid has potential for cross-sectional performance evaluation in early growth stages, but that power wanes as the model shifts its focus to the unique cultures. This does not necessarily diminish the potential usefulness of the model. However, the degree of generalizability does affect how the model should be used.

Delineating Granularity and Frequency

Usefulness of a model is also contingent on knowing when and where to apply the model based on the inherent temporal characteristics of the model and the accompanying data. The Pyramid as presented in Flamholtz (2005) could be considered a broad, episodic performance measurement model. In a growing firm, the framework is designed to detect the need for three significant changes, one each time a company exceeds the limits of a growth stage.[2] From the information presented (Flamholtz, 2005), the model does not appear to be designed to detect small changes over time affecting performance within a growth stage. Other performance measurement models, such as the balanced scorecard (Kaplan & Norton, 1996, 2001) and the performance pyramid (Lynch & Cross, 1991), have a greater potential to detect small changes that affect performance throughout a company's life. These two models can be characterized as more detailed, continuous use models. In an other work, Flamholtz (2003) suggests that the Pyramid can also be used in a continuous fashion by using the six building blocks as performance measurement model categories instead of the balanced scorecard's four perspectives. However, Flamholtz and his co-authors have not recommended or tested specific qualitative or quantitative measures within each of the six building blocks which would facilitate its use as a continuous model.

MODEL TESTING

Our interpretation of the overall Pyramid of Organizational Development framework is given in Fig. 1.

Any proposed holistic performance measurement model needs to be supported by well-designed and executed research. To be sure, the process of

Fig. 1. Pyramid of Organizational Development Framework.

model testing and validation is lengthy and tedious. Flamholtz refers to six published studies, all published in the *European Management Journal*, to demonstrate what he identifies as at least preliminary results to support the model. Four are based on multiple divisions of the same company (Flamholtz, 2001; Flamholtz & Hua, 2002a, b; Flamholtz & Kannan-Narasimhan, 2005) and two are based on the same set of companies (Flamholtz & Aksehirli, 2000; Flamholtz & Hua, 2003). Although Flamholtz (2005) does not indicate so, the samples drawn seem to be theoretical samples (Glaser & Strauss, 1970) chosen to help build the model, which is an appropriate research strategy during model development. Model testing requires additional sampling strategies.

A major implication from the framework is that there must be fit between the degree of organizational development (i.e., the six building blocks) and stage of company growth (Fig. 1, Box A). If fit is not achieved, then growing pains will result (Fig. 1, Box B), leading to poor financial performance (Fig. 1, Box C). None of the six empirical studies outlined in Flamholtz (2005) investigate the first link in the model. One published study, Flamholtz and Hua (2002b), tested the relation between growing pains and financial performance. Since growing pains and financial performance are measured contemporaneously, it is difficult to determine if the growing pains actually preceded the poor financial performance. Neither human resource researchers nor empirical results are in agreement about whether employee attitudes, which Flamholtz's growing pains appear to reflect, influence business outcomes or whether business outcomes influence employee attitudes (Koys, 2001). Once again, circularity comes into question.

Another, more basic implication from the model is that the six building blocks are drivers of financial performance (see Fig. 2). The majority of the empirical tests (Flamholtz & Aksehirli, 2000; Flamholtz, 2001; Flamholtz & Hua, 2002a, 2003; Flamholtz & Kannan-Narasimhan, 2005) examine this link between the degree of organizational development and financial performance, regardless of fit with growth stage. The results of the Pyramid studies are encouraging, in that the Pyramid appears to capture relevant determinants of financial performance. Flamholtz is making some initial

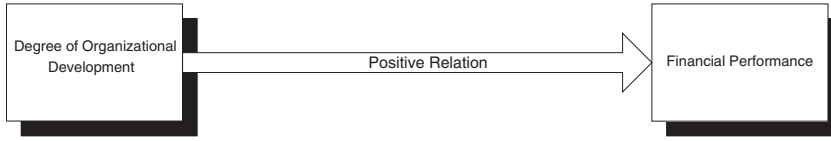


Fig. 2. Building Blocks as Performance Drivers.

efforts at providing evidence to support the model. However, at this point the results are, as Flamholtz (2005) clearly states, preliminary.

IMPLICATIONS OF EMPIRICAL RESULTS

As argued previously, a well-developed holistic performance measurement model should be useful for both analytic and predictive purposes. The research necessary to support such use of a model will of necessity involve, as discussed by Flamholtz (2005), longitudinal studies. The promising findings of correlations among some pieces of the model are a first step toward establishing causality. Temporal precedence needs to be established (e.g., growing pains occurring before financial performance suffers) before a claim of predictive ability can be made. Given the research cited in Flamholtz (2005), it would seem more appropriate to limit any use of the model to classification rather than prediction. At this stage of model development, empirical testing to date is useful for generalizing to theory, rather than generalizing to a population (Yin, 1994). In the future, other researchers can test this framework with randomly selected companies in order to generalize their results to populations of firms.

CONCLUDING COMMENTS

As the trite old expression goes, ''behold the turtle, he only makes progress by sticking his neck out.'' This is not to imply our colleague is a turtle or turtle like but he has and does stick his neck out. In doing so, he delivers a foundation that can help other researchers develop better models. For instance, today it is very popular to be a researcher studying intangibles. Flamholtz was attempting to deal with intangibles long before it was popular. His work in human resources accounting in the 1970s (e.g., Flamholtz, 1971) was one of the early serious attempts at the analysis of intangibles.

We are able to criticize his model from various perspectives primarily because he has held it up to be critiqued. More important than our comments is that he is developing a model for us to critique.

NOTES

1. We have attempted to limit our discussion to the model as presented in Flamholtz (2005). To help make this paper coherent, we did at times find it necessary to refer to other published formulations of the Pyramid. We have used the published articles referenced in Flamholtz (2005) to discuss the development and testing of the Pyramid.
2. Flamholtz (2005) and Flamholtz and Hua (2002a) present four stages of growth while Flamholtz and Randle (2005) presents seven stages.

REFERENCES

Flamholtz, E. (1971). A model for human resource valuation: A stochastic process with service rewards. *The Accounting Review*, *46*(2), 253–267.

Flamholtz, E. (2001). Corporate culture and the bottom line. *European Management Journal*, *19*(3), 268–275.

Flamholtz, E. (2003). Putting balance and validity into the balanced scorecard. *Journal of Human Resource Costing and Accounting*, *7*(3), 15–26.

Flamholtz, E. (2005). Strategic organizational development and financial performance: Implications for accounting, information, and control. *Advances in Managerial Accounting*, *14*, 141–168.

Flamholtz, E., & Aksehirli, Z. (2000). Organizational success and failure: An empirical test of a holistic model. *European Management Journal*, *18*(5), 488–498.

Flamholtz, E., & Hua, W. (2002a). Strategic organizational development and the bottom line: Further empirical evidence. *European Management Journal*, *20*(1), 72–81.

Flamholtz, E., & Hua, W. (2002b). Strategic organizational development, growing pains and corporate financial performance: An empirical test. *European Management Journal*, *20*(5), 527–536.

Flamholtz, E., & Hua, W. (2003). Searching for competitive advantage in the black box. *European Management Journal*, *21*(2), 222–236.

Flamholtz, E., & Kannan-Narasimhan, R. (2005). Differential impact of cultural elements on financial performance: An empirical investigation. *European Management Journal*, *23*(1), 50–75.

Flamholtz, E., & Randle, Y. (2005). *Measuring and managing successful organizational scale-up forthcoming in conceptualizing scale-up: Multidisciplinary perspectives*. Chicago, IL: The University of Chicago Press,.

Glaser, B. G., & Strauss, A. L. (1970). Theoretical sampling. In: N. K. Denzin (Ed.), *Sociological methods* (pp. 105–114). Chicago: Aldine.

Greiner, L. E. (1998). Evolution and revolution as organizations grow. *Harvard Business Review*, *76*(3), 55–67.

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard*. Boston, MA: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. (2001). *The strategy-focused organization*. Boston, MA: Harvard Business School Press.

Koys, D. J. (2001). The effects of employee satisfaction, organizational citizenship behavior, and turnover on organizational effectiveness: A unit-level, longitudinal study. *Personnel Psychology*, *54*(1), 101–115.

Lynch, R. L., & Cross, K. F. (1991). *Measure up! Yardsticks for continuous improvements*. Oxford, UK: Blackwell.

Simons, R. (1995). *Levers of control*. Boston, MA: Harvard Business School Press.

Yin, R. K. (1994). *Case study research design and methods*. Newbury Park, CA: Sage.

This page intentionally left blank

THE PYRAMID OF ORGANIZATIONAL DEVELOPMENT AS A PERFORMANCE MANAGEMENT AND MEASUREMENT MODEL: A REPLY

Eric G. Flamholtz

Euske and Malina (2005) have presented a thoughtful and constructive critique of my article entitled "Strategic Organizational Development and Financial Performance: Implications for Accounting, Information, and Control." However, I disagree with a number of questions and criticisms they have raised.

PERFORMANCE MANAGEMENT VERSUS PERFORMANCE MEASUREMENT

Euske and Malina (2005) view the article as a "performance measurement model." Actually, it is intended as an organizational development/performance management model with implications for performance measurement as well as other managerial and accounting applications. This is not just a semantic quibble; it is central to understanding the article and the underlying stream of theoretical and empirical research that it reports.

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 177–188

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14008-8

My primary interest with The Pyramid of Organizational Development model is in providing a holistic model for building successful organizations. It is a model of organizational performance creation, which can be used as a model for performance measurement.

The companion life cycle model is intended to address the issue of the need for organizational change and transition over time as a result of changes in either the environment or size of the organization (Flamholtz, 1995; Flamholtz & Randle, 1998, 2000). Another companion model is of the organizational "growing pains" which are hypothesized to emerge because of a lack of an effective fit between the development of the Pyramid and the given stage of growth of an organization, as measured in terms of "size" with (as discussed below) revenues used as a surrogate measure of organizational size.

Taken together, these three models comprise the overall theoretical framework that I have developed. All of the theoretical and empirical work surveyed in Flamholtz (2005), current volume is intended to address this overarching issue of building successful organizations at different stages of growth.

Clearly, once we have a holistic framework for performance management, the logical application is as a model for performance measurement; but the latter is derived from the former. This distinction is critical because the requirements, perspective, and criteria for testing the performance management model are somewhat different from that of testing a performance measurement model. For example, accounting is a system for measuring economic events and transactions that are, in part at least, of an economic character. The criteria for accounting systems are not the same as for modeling the underlying economic activity to which the accounting measurements are addressed. This suggests that the perspective used by Euske and Malina, while appropriate to performance measurement, might not be equally appropriate to organizational development and performance management. Accordingly, as I shall show below, some of the comments and critique presented by Euske and Malina (2005) are inappropriate (lack validity) because of their measurement perspective on my models and research. Nevertheless, I will address their specific comments and questions. In general, I will organize this reply according to the sections of their paper.

MODEL CHARACTERISTICS

This is an area where Euske and Malina's measurement perspective provides too narrow a look at the organizational development/performance management model. However, I will still respond to the measurement issues per se.

Euske and Malina (2005) begin with the assertion that "if a comprehensive performance model for business is to be useful in an analytic and predictive sense, the model must capture the interrelationships of factors that influence organizational performance such as organizational; maturity, size, products and services, management systems, industry characteristics and environmental influences." What they fail to recognize is that their assertion is, in fact a testable hypothesis, not a statement of empirical fact. Implicitly, they are proposing an alternative model, which, then, of course, would require empirical testing. Indeed, this is exactly what the Pyramid of Organizational Development is designed to address. If Euske and Malina wish to Test their alternative model of organizational performance, I wish them well.

Revenues as a Measure of Growth Stages

Euske and Malina (2005) raise interesting questions about the use of revenues to identify and define growth stages. These issues were beyond the scope of the original paper, but they are relevant issues.

There are two key aspects of their questions. One of them concerns the use of levels of sales revenue with specific growth stages.

Based upon my analysis and experience with organizations, I have reached the conclusion that sales revenue should be used as a surrogate or proxy measure of organizational size, and therefore to define stages of growth. I recognize that this is a testable hypothesis, and I have not yet tested it. However, I have concluded that revenues are superior to other possible alternatives because of greater "face validity." Other possible measures that have been used in the life cycle literature include organizational age (chronological), number of employees, and "undefined" size. Why is revenue the best measure on a face validity basis? Organizations can pass through various stages of growth from "birth" to "death" (bankrupt and/or "out of business") independent of age. Both Kemper insurance and Osborne Computer have ceased to exist; but Kemper was more than 70 years old while Osborne Computer was 3 years old. The number of employees is a judgmental thing; not the counting of employees, but the number that is on hire. Sales revenue (except at places like Enron!) is a relatively "objective" measure, even given all of the known limitations of revenue recognition in GAAP.

Euske and Malina (2005) cite Greiner (1998), who uses organizational age to define stages of growth, as implicitly presenting a superior model. I challenge that assertion, and in fact I believe that one of the advantages of

my model is it can be used because of the greater face validity of sales revenue rather than age, as used in the Greiner (1998) model.

Euske and Malina (2005) also raise the issue of whether size as measured by revenue is an appropriate measure because of possible exceptions. As they state: "A manufacturing organization, such as a shipyard, could be a relatively new venture with only one order and be in the highest sales revenue category of the Flamholtz model." They are correct, but that is not the whole story. This would be an anomaly and when it occurs it would be problematic for the organization, and a likely predictor of growing pains, as explained below.[1]

Although this issue was not addressed explicitly in the current paper because of scope issues (Flamholtz, 2005), it has been addressed in prior publications that were cited as support for the models (Flamholtz & Randle, 2000). Specifically, Flamholtz and Randle (2000, p. 29) state: "The stages of organizational growth, the critical tasks of organizational development for each stage, and the approximate size (measured in millions of dollars of sales revenues) at which an organization will typically pass through each stage are shown in Table 2.1.[2] A key word in this statement is typically. What this means is that for approximately 90% of manufacturing firms that have revenues in the range of $10 million to $100 million, they will typically encounter the critical issues of stage III. However, some firms will have to face these problems at a smaller size in development or much later in their development." We also stated (Flamholtz & Randle, 2000): "Accordingly, we need to view the relevant range as designated for the transitions to occur at each stage of development as a 'normal curve'."

Although in hindsight it would have been better to add these statements to the current paper for clarification, it is not possible to summarize an entire book in a single article and it was incorporated by reference.

Assessing Issues of Use versus Design

In this section, Euske and Malina (2005) raise the question of whether growing pains are the result of the design of the six key building blocks comprising the pyramid or their use (misuse?), or both. With respect to their question, it does not matter except in a very narrow sense; that is, the distinction has no practical significance. The model specifies the degree of distress caused when there is not a good fit between the size of the organization and its infrastructure. Just like pain in the human body, it is a "warning" that something is wrong, and a call to action to fix it. Obviously, further analysis is required to identify the source of the problem. What the growing pains measure does is to identify that there is a problem caused by

"inadequate infrastructure, regardless of whether its cause is lack of development or problems in use. My experience in this regard has indicated that the typical cause is underdevelopment of infrastructure, but it is theoretically possible to be problems in use. There is a companion survey (The Survey of Organizational Effectiveness), which is a validated 65-item Likert Scale-based instrument, which helps to pinpoint the specific problems and causes of organizational distress. I have been using this for several years for research and consulting practice. In addition (as discussed below), there is empirical evidence to support my position on this issue.

Intent of the Models

Euske and Malina (2005) also state that "the logic behind the model appears to be that it is designed for relatively large growing organizations that render the infrastructure of the organization obsolete at regular intervals." Not quite; the models (note the plural) are designed as life cycle models for organizational success at different stages of growth, from a new venture through organizational maturity and then decline.

Specifying Causal Relationships

Euske and Malina (2005) state "…in Flamholtz model, the assumption that the misfit between sales revenue and organizational infrastructure causes growing pains is tenuous." This is not quite what I have said. I have said that the misfit between **organizational size and complexity, as measured by sales revenue as a surrogate or proxy measure, causes growing pains**. This, of course, is a testable hypothesis, but it has a great deal of face validity. As discussed below in the section titled "Pyramid Development and Growing Pains: New Data," previously unpublished data from the Banner corporation studies (Flamholtz & Hua, 2002a, b) indicate that there is a statistically significant relationship between the degree of development of the pyramid and growing pains ($r^2 = 0.40$, significance = 0.01).

Defining the Degree of Generalizability

This is an interesting issue raised by Euske and Malina (2005), and I agree with the general thrust of their argument. However, a few key points ought to be noted: The pyramid framework provides the very kind of generalizability sought by Euske and Malina. It argues that these are the key phases of the "business game." With respect to the measurement of the six key phases, it is correct that six questions were used in Flamholtz and Hua (2002a) to assess the level of organizational development of the Pyramid. However, the

validated Survey of Organizational Effectiveness (cited above) which is a 65-item instrument was used in Flamholtz and Kurland (2005), and will be used in the future. This is a complex field and progress in measurement and cumulative research occurs step by step, as Euske and Malina have noted.

Detailing Granularity and Frequency

This is another area where Euske and Malina's measurement perspective is too narrow a look at the organizational development/performance management model.

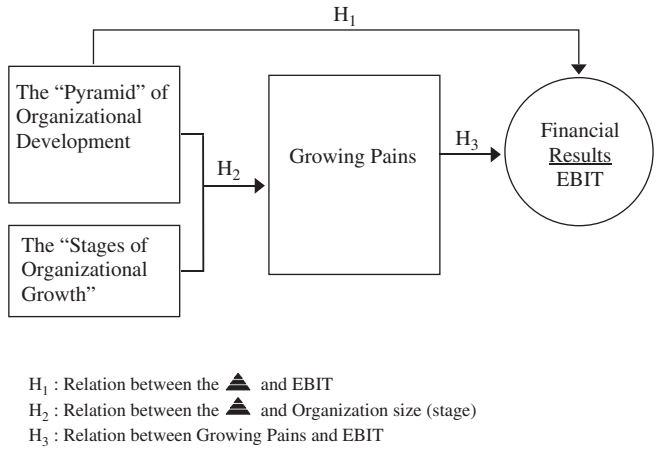
Euske and Malina (2005) state that "other performance measurement models, such as the balanced scorecard (Kaplan & Norton, 1996, 2001) and the performance pyramid (Lynch & Cross, 1991), have a greater potential to detect small changes that affect performance throughout a company's life." Although the balanced scorecard is widely promoted and has been applied, it has not been supported by the kinds of empirical research cited in my paper for the pyramid (Flamholtz, 2003). In addition, there are many applications of the pyramid as what Euske and Malina have termed "continuous use models." It has been used as part of an overall strategic management system for more than 20 years by many companies ranging from large Fortune 200 companies such as Countrywide Financial Corporation, Starbucks, and PacifiCare to small new ventures and midsized companies. Some of these applications have been described in Flamholtz and Randle (2000). Others will be described in more depth in a book in progress on leading strategic and organizational change.[3]

Euske and Malina (2005) also state that: "however, Flamholtz and his co-author have not recommended or tested qualitative or quantitative measures within each of the six key building blocks which would facilitate its use as a continuous model." This statement is not unreasonable, but it is not totally correct. As noted above, the Survey of Organizational Effectiveness, which is a 65-item instrument was used in Flamholtz and Kurland (2005), and has been used in many other applications. However, this is a proprietary instrument and it has not been placed into the public domain.

MODEL TESTING: CRITERIA, DIFFICULTIES, AND PROGRESS

This section has a number of issues that need to be addressed. First, the schematic representation of the theoretical framework provided by Euske and Malina (2005) is a reasonable interpretation of the overall set of models.

Exhibit 1. Graphic Summary of the Research Framework.



However, it is not quite correct. The schematic representation of the framework as I use it is shown in Exhibit 1.

As seen in Exhibit 1, there is a hypothesized direct relationship between development of the Pyramid and financial performance as well as a relationship between the fit between the pyramid and stages of growth to growing pains and financial performance. As can be seen, this is a complex set of proposed relationships, and since 1995, I have been engaged in a series of empirical studies to test the hypotheses.

As Euske and Malina state, there have been six published articles (during the period 2001 to 2005) presenting empirical research to test aspects of the theoretical framework. There is also one additional paper (Flamholtz & Kurland, 2005) that is forthcoming. As Euske and Malina point out, four are based on data from multiple divisions in one company and two from another (cross sectional) set of companies. I do not know whether their description of this fact was intended as a criticism or merely description. But here is the thing: empirical research is difficult and time consuming. It is difficult to get data from companies. They typically require some quid pro quo. I learned that from working as a Research Assistant for the late Rensis Likert, many moons ago.

I waited for more than 10 years to get access to a research site of the kind found in ''Banner corporation,'' which is the company cited in the four

studies using comparable divisions. The forthcoming study by Flamholtz and Kurland (2005) comes from another type of business, a third data set, also requiring more than one year to collect the data.

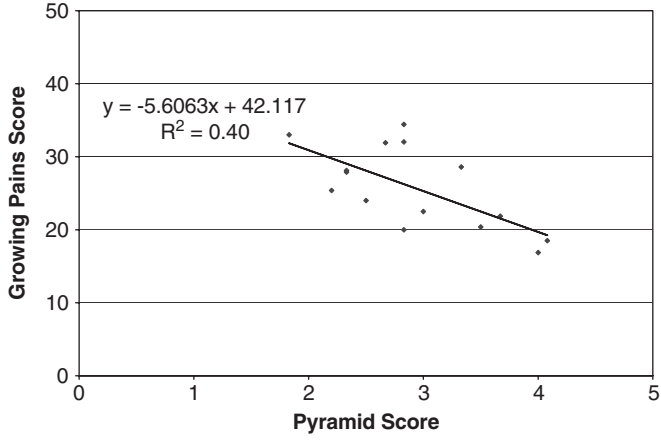
The data for the other two published articles took more than 1 year to collect and analyze. There is no set of published tapes with the data required for this type of research.

Obviously I agree with the statement that: ''any proposed holistic performance measurement model needs to be supported by well designed and executed research.'' The lack of this research in support of the so-called ''balanced scorecard'' is one of my primary criticisms of that construct (Flamholtz, 2003). I have published six empirical studies testing the models since 2000, and another is forthcoming. This is a substantial body of research for this time period.

Pyramid Development and Growing Pains: New Data

Interestingly, Euske and Malina cite an omission in published research concerning the hypothesized relations between strategic organizational development (the pyramid) and growing pains. They are correct about the omission, but there is substantial unpublished data about the predicted

Exhibit 2. Pyramid Summary of the Research Framework.



relationship.[4] Data on the relationship between strategic organizational, development of the pyramid and growing pains for Banner Corporation (the company used as the research site in four articles based upon data in a single organization) is shown in Exhibit 2 R^2 for these data is 0.40 and is statistically significant at less than 0.01 (0.007). This provides strong support for the predicted relationship between the pyramid and growing pains. This relationship has been replicated many times in other as yet unpublished studies. However, Euske and Malina were correct to point out this omission.[5]

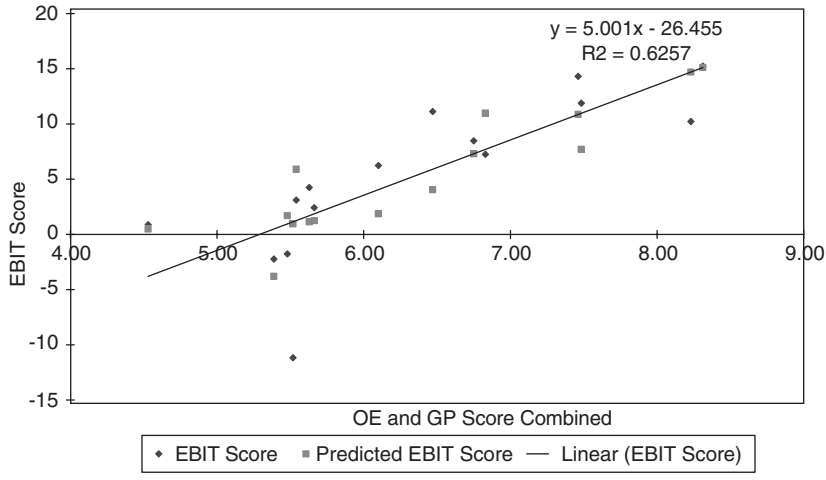
IMPLICATIONS OF EMPIRICAL RESULTS

Euske and Malina (2005) state "Given the research cited in Flamholtz (2005), it would seem more appropriate to limit any use of the model to classification rather than prediction." They also state: "In the future, other researchers can test this framework with randomly selected companies in order to generalize their results to populations of firms." I cannot disagree more strongly with both assertions. First, the data presented in the empirical studies indicate the predictive value of the models. They are potentially an important supplement to conventional financial information because they are ex ante data, not ex post data like conventional accounting information. Second, the idea of studying this type of model with a random sample sounds sensible but is spurious. The explanatory power of the empirical studies is that they have minimized "noise" from extraneous factors to the maximum extent feasible. Instead of random samples, which give a spurious sense of scientific rigor, we need samples that eliminate as many extraneous factors as possible to isolate the control and experimental effects. That is why I chose a "paired comparison design" from Flamholtz and Aksehirili (2000), and to study multiple divisions within the same firm, where profit centers are approximately the same size as in Flamholtz (2001), Flamholtz and Hua (2002a, b), and Flamholtz and Narasimhan (2005).

USE IN MANAGEMENT CONTROL

There is an additional potential use of this framework and the related measurements, which was not discussed in the original paper. Assuming that internal managerial accounting is defined broadly to include internal control there is another potential application and use of the proposed models and related measurements. We have been doing research on the use of the

Exhibit 3. Organizational Risk Assessment Score- (Organizational Effectiveness and Growing Pains scores combined) Significance = 0.0045.



combined measures of organizational development and growing pains as a surrogate measure of ''operational risk'' in context of Sarbanes–Oxley. Using data from the banner Corporation studies (Flamholtz & Hua, 2002a, b) we have found that the combined measure of growing pains and organizational development explains about 63% of Earnings Before Interest and Taxes (EBIT), as shown in Exhibit 3. This measure is statistically significant at an extraordinarily high level (0.00045).[6]

This is potentially powerful too to measure operational risk. It can be used by external auditors and the Board as well as by internal management accountants concerned with operational control.

THE METAPHOR OF THE TURTLE

Euske and Malina (2005) conclude with the metaphor of a Turtle sticking its neck out. They are gracious and complementary in the use of this metaphor, and I was not offended at all. Although I might have initially preferred the metaphor of a Lion or a Bear, the turtle too has a noble history! The ancient

Romans used the formation of a Turtle to great tactical advantage. The sea turtle swims among sharks, including Great White Sharks, and heads out into the vast unknown of the perilous ocean. The turtle also defeated the rabbit in the fabled race. So I love the metaphor!

Much of the literature in my article had to be new to Euske and Malina, and I and the editors agree that they did a very nice job in raising issues and making critical comments. Although I do not agree with much of their analysis, nevertheless, it was done in a constructive way with an appropriate "voice."

CONCLUSION

I hope the original paper as well as their critique and my reply are of interest. I would welcome others to this "ocean." I believe that the theoretical framework as well as the related empirical research presented in the original article has the potential of a major paradigm shift not only in management but in accounting as well. I am pleased to share it with my colleagues in accounting. My sincere thanks to Euske and Malina for their time, efforts, and thoughtful feedback; this is what scholarship is supposed to be. Finally, welcome to the Ocean. It gets lonely being the lone turtle in a sea of sharks!

NOTES

1. This is, in fact, why Osborne Computer went bankrupt, because there was a lack of fit between its size and its infrastructure.
2. This is the same as Exhibit 2 in Flamholtz (2005).
3. This book is under contract to Cambridge University Press and is expected to be published in 2007.
4. Frankly, I am not sure why I have not published it except that it has been so obvious to me for so long that I have taken it as a given. I have been more concerned about the relation between the pyramid and financial performance and between growing pains and financial performance.
5. I am grateful to Euske and Malina for identifying this (unconscious but obvious) omission and am pleased to add this data to the growing body of empirical research designed to test the models.
6. There is virtually a zero probability (significantly less than 1%) that this could occur by chance.

REFERENCES

Flamholtz, E. (1995). Managing organizational transitions: Implications for corporate and human resource management. *European Management Journal*, *13*(1), 39–51.

Flamholtz, E. (2001). Corporate culture and the bottom line. *European Management Journal*, *19*(3), 268–275.

Flamholtz, E. (2003). Putting validity and balance into the balanced scorecard. *Journal of Human Resource Costing & Accounting*, *7*(3), 15–26.

Flamholtz, E. G., & Aksehirli, Z. (2000). Organizational success and failure: An empirical test of a holistic model. *European Management Journal*, *18*(5), 488–498.

Flamholtz, E., & Hua, W. (2002a). Strategic organizational development and the bottom line: Further empirical evidence. *European Management Journal*, *20*(1), 72–81.

Flamholtz, E., & Hua, W. (2002b). Strategic organizational development, growing pains and corporate financial performance: An empirical test. *European Management Journal*, *20*(5), 527–536.

Flamholtz, E., & Kannan-Narasimhan, R. (2005). Differential impact of cultural elements on financial performance: An empirical investigation. *European Management Journal*, *23*(1), 50–64.

Flamholtz, E., & Kurland, S. (2005). Strategic organizational development, infrastructure, and financial performance: An empirical investigation. Unpublished working paper to be presented at the UCLA/ Oxford University conference on research in entrepreneurship, June 22–24.

Flamholtz, E. G., & Randle, Y. (1998). *Changing the game: Transformations of the first, second and third kinds*. New York, NY: Oxford University Press.

Flamholtz, E. G., & Randle, Y. (2000). *Growing pains: Transitioning from an entrepreneurship to a professionally managed firm*. San Francisco, CA: Jossey-Bass Publishers Inc.

Greiner, L. E. (1998). Evolution and revolution as organizations grow. *Harvard Business Review*, *76*(3), 55–67.

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard*. Boston, MA: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. (2001). *The strategy-focused organization*. Boston, MA: Harvard Business School Press.

Lynch, R., & Cross, K. (1991). *Measure Up! Yardsticks for continuous improvement*. Oxford, UK: Blackwell Publishers.

EARLY EVIDENCE ON THE INTERACTIVE EFFECTS INVOLVING PRODUCT DEVELOPMENT ORGANIZATIONS AND TARGET COST MANAGEMENT

Chao-Hsiung Lee, John Y. Lee and Yasuhiro Monden

ABSTRACT

In this article, we examine the link between product development organization and target cost management. More specifically, we investigate the interactive effects of alternative product development organizations, methods for setting target costs, and alternative decision-making authority in assigning targets. Based on the results of a questionnaire survey of Japanese manufacturers, we intend to provide some early evidence on those interactive effects to stimulate further research in this area of target cost management. We find that organizational efficiency is connected with cost reduction performance in target cost management from the two perspectives: (1) the relationship between the simultaneous involvement of project leaders and functional staff and the use of a particular target cost-setting method, and (2) the

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 189–209

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14009-X

correlation of the simultaneous project-function involvement with the level of expected cost reduction performance. No particular product development organization shows any preference for the use of a target cost-setting method. The preference, however, is shown clearly when a particular target cost allocation decision authority is aligned with the use of a target cost-setting method. Companies with higher levels of cost reduction performance (expected or experienced) tend to make project leaders and functional staff get simultaneously involved in target cost allocation decisions.

INTRODUCTION

Target cost management, a strategic cost-management technique, is used by many manufacturing and service firms in the world. An empirical test has even elevated its use to a facilitator of organizational learning in an advanced technology environment (Choe, 2002). Among the successful users of target cost management, Toyota's case has been well publicized thus far (Ansari, Bell, & the Target Cost Core Group, 1997; Cooper & Slagmulder, 1997). Toyota has let the Japanese auto industry know that its success in target cost management depends heavily on its efficient product development organization. Toyota's two types (matrix and multi-center) of product development organizations have become very popular among Japanese automakers and manufacturers in other industries as well.[1]

In this study, we examine the link between product development organization and target cost management. More specifically, we investigate how product development organizational structure, as an organizational alignment (Clark & Fujimoto, 1991; Powell, 1992; Kloot, 1997), is related to the target costing method selection and target allocation decisions in the process of target cost management. This study seeks to add to the target cost management literature by analyzing the interactive effects of alternative product development organizations, methods for setting target costs, and alternative decision-making authority in assigning targets. We intend to provide some early evidence on those interactive effects to stimulate further research in this area of target cost management.

The remainder of this paper is organized as follows: In the next section, we present the target costing framework and develop the hypotheses to be tested. Then we explain the questionnaire survey we have performed and analyze the results. Finally, we discuss future research implications of our findings.

FRAMEWORK AND HYPOTHESES

Product Development Organization

Toyota has introduced the two types (matrix and multi-center) of product development organization first. The two types of product development organization have become popular among Japanese manufacturers (see the survey mentioned in Note 1).

In a matrix organization of product development, product project department and functional design department are separately established. In each product project department, product project leaders are responsible for conceptualizing a product, drafting the basic product plan, and establishing the product target cost. The functional staff designs the functional parts of the product. Using the matrix structure (see Fig. 1), project leaders assigned to specific models (or products) work with the functional staff assigned to specific functions related to the model or product.

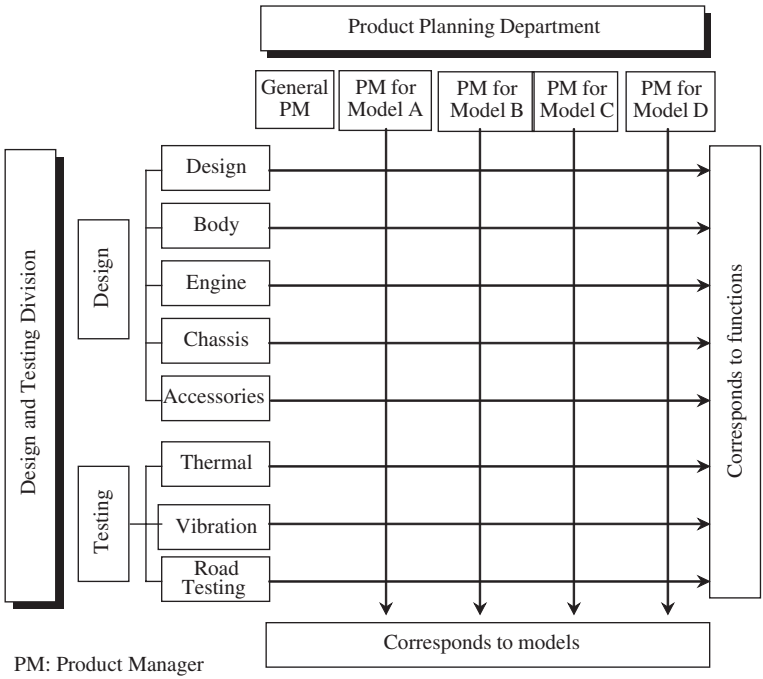


Fig. 1. Matrix Organization of Product Development.

Product project department and functional design department are established as independent and separate units in a matrix organization and a project leader is a project department manager who takes responsibility for a new product project (See Model A or Model B in Fig. 1). A functional manager is the leader of a functional design department, and the functional staff involved in new product development reports to both the project leader and the functional manager.

In a multi-center organization of new product development, several project centers may be established in the organization and a functional design group is created in each center (Cusumano & Nobeoka, 1998); Fig. 2 shows the structure. Each center develops a new product while coordinating with its own functional design group. Toyota's functional design group (not shown in Fig. 2), for example, assists with plans on the design and test involving style, body, interior, chassis, and engine. The functional staff at each center reports to the center head.

All products are assigned to respective product development centers, according to the nature of the products, in a multi-center organization. All projects in a center are developed using a common platform. Toyota

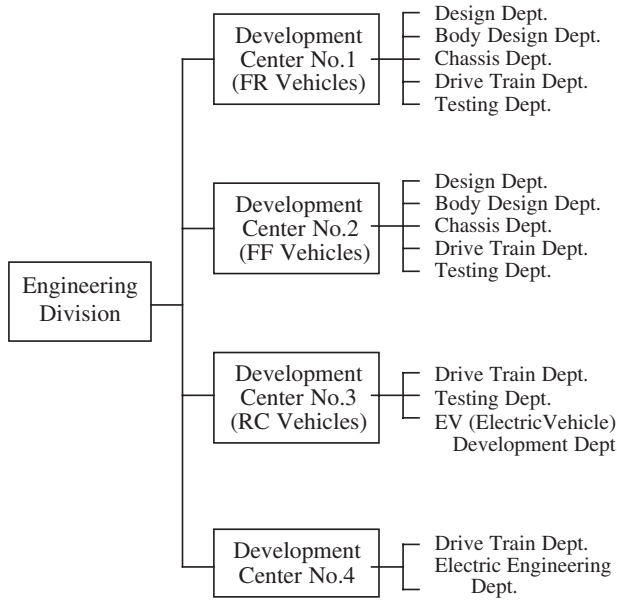


Fig. 2. Multi-Center Organization of Product Development at Toyota.

developed the multi-center organization in 1992. Prior to 1992, Toyota used the matrix organization of new product development. Toyota's problems with the matrix organization included poor communication between project leaders and functional staff and the long lead time in decision-making on product design and development. The great variety of products created additional problems in the allocation of capital and personnel, the transformation of production technology, and the adoption of common parts. The multi-center organization of product development was adopted as a more flexible and less sophisticated version (Cusumano & Nobeoka, 1998).

Methods Used to Set Targets

Setting the target cost of a new product is the most critical step in a firm's target cost-management process, and is affected by the firm's product strategy and long-term profit planning (Ansari et al., 1997). Seventy to ninety percent of all costs are typically determined during the product development stage and it is important to set an attainable target cost (Euske, Lebas, & McNair, 1993; Koga & Matsuo, 1996; Ansari et al., 1997). Japanese companies set target costs in three different ways: subtraction, hybrid, and addition (Tanaka, 1989; Kondo, 1990; Kanazawa & Monden, 1994).

Subtraction Method

Under the subtraction method, target sales price and provisional target profit are first determined. Target cost is calculated by subtracting target profit from target sales price or by multiplying "1 – target return-on-sales ratio" by target sales price. The "allowable cost," as measured by the right-hand side of Eq. (1), is treated as the target cost:

$$\text{target cost} = \text{target sales price} - \text{target operating profit}, \quad (1)$$

Or

$$\text{target cost} = \text{target sales price} \times (1 - \text{target return-on-sales ratio}) \quad (2)$$

This method reflects a customer-oriented concept: it is based on such factors as the number of products to appear on the market, the price of rival products, and the level of desired profit. The method incorporates a

particular development concept: "How can we develop a product that is both good for us and easily recognizable by our customers?" Companies that use this method aim to manage product costs effectively while coping with market pressure (Kondo, 1990), as Fig. 3 indicates.

To deal with the market pressure, companies are motivated to improve technology and production management, and reinforce product innovation. Since *future* technology level is factored in the determination of target cost, the target cost is not easily attainable in actual practice. Estimated cost based on the current technology is not the level considered, and thus, target cost is tightly set under the subtraction method.

Addition Method

Allowable costs are determined under the relevant market conditions without any allowance for the design capability or production technology of the firm. As a result, the resulting cost reduction targets are very stringent (Monden & Hamada, 1991). Companies whose allowable costs are beyond any reach can consider the use of estimated costs. (Fig. 4 shows the flow). Under the addition method, estimated cost plus the provisional target profit will determine the value of target price. Target cost is calculated by subtracting per-unit profit improvement target from the estimated cost, assuming the use of the present technology by the firm. The estimated cost of a product is the "addition" of the estimated costs of all parts used in the product, as shown in the following equation:

$$\text{target cost} = \sum \text{estimated cost } i - \text{profit improvement target} \quad (3)$$

where profit improvement target = estimatedprofit − targetprofit.

Fig. 3. Customer-Based Target Cost Setting.

Fig. 4. Technology-Based Target Cost Setting.

This is a traditional method in the sense that a target cost and a target sales profit will be used to determine the target sales price. The addition method is technology-oriented and those companies that use this method value the importance of developing new production technology to make a good-quality, low-priced product (Kondo, 1990).

Hybrid Method

Some companies consider both allowable and estimated cost in the computation of target cost. Under this hybrid method, the resulting target cost usually reflects the two cost figures as

$$\text{target cost} = w_1 S + w_2 A \tag{4}$$

where S is the target cost under the subtraction method, A the target cost under the addition method, and

$$w_1 + w_2 = 1$$

Hypothesis on Product Development Organization and Target Cost-Setting Method

Each functional staff usually supports three or four product projects in a matrix organization. This practice helps maintain the efficiency of product development, and small-size companies prefer the functional matrix organization (Cusumano & Nobeoka, 1998). Matrix organizations enjoy the economy of scale in terms of information processing and technology application by sharing the pools of functional staff assigned to support product development (Davis & Lawrence, 1977; Shibata & Nakahashi, 1997).

Both vertical and horizontal authority relationships exist in matrix organizations. The functional staff reports to the respective project leader, and is responsible to the functional manager also. Each product-development project team is a temporary organization. The functional staff recruited from different functional departments work together on project teams. Both project leaders and functional managers have the authority and the responsibility for product development. However, a functional staff usually tends to identify with one's own functional manager when conflicts arise between the project leader and the functional manager. A functional convenience is usually preferred to a project success in the workplace, favoring a functional rather than a marketing focus in setting target costs. Accordingly, a matrix organization would tend to favor the addition method rather than the subtraction method. Target sales price is market-based and

project leaders tend to fix target sales profit, and no compromise in setting cost targets is expected under the subtraction method.

In a multi-center product development organization, each project leader must establish a cost budget for individual products, requiring project leaders to set cost targets and break component costs down to product parts in consultation with the functional staff. Also, each product development center, established for market segmentation and customer support, develops and produces selected products for special-class customers. The market focus and the simultaneous involvement of product managers and the functional staff would make the subtraction method look more compatible with a multi-center organization.

Accordingly, we hypothesize the following:

H1. Companies with a multi-center product development organization are more likely to use the subtraction method of target cost determination.

Hypothesis on the Target Cost-Setting Method and the Target Cost Allocation Authority

A multi-center organization of product development allows each center to be determined based on the attributes of product development projects. Each center develops the products with compatible attributes on a common platform, and project leaders work with the functional design teams that are assigned to each center. Each project leader is responsible for one product development project and the assigned functional design staff reports to the project leader. Project leaders are empowered by top management through the entire process of product development and cost reduction.

Product managers control target cost allocations and the allocations are based on cost elements and functional specifics. Cost element-based allocations rely on the attributes of costs (materials, labor, depreciation, etc.). Function-specific allocations are related to the vehicle's structural components corresponding to the various functional departments. The costs are allocated to functional specifics corresponding to relevant functional departments and each functional department deals with function- or part-specific target costs.

Since (1) the subtraction method is the most market oriented, as discussed in the previous section, among the three methods, and (2) each center in the multi-center organization (market oriented) has both product manager and the functional staff, we hypothesize a probable relationship between the use of a subtraction method and the simultaneous involvement of the product

manager and the functional staff in target cost allocation decisions. This hypothesis would help us find whether the target cost allocation authority is aligned with the selection of a particular target cost-setting method. The formal hypothesis is formulated as follows:

H2. Companies with a market orientation strategy would prefer a close alignment of the target cost allocation authority with the subtraction method.

Hypothesis on Target Cost Allocation Authority and the Cost Reduction Performance Level

Would a company that requires a simultaneous involvement of the product leader and the functional staff in target cost allocation decisions shows a higher level of cost reduction performance? According to Cusumoto, Nonaka, and Nagata (1995), the exchange of knowledge among teams and the interaction among staffs produce a dynamic communication that increases interdepartmental coordination and communication.

The interaction between the project leader and the functional staff and the resulting knowledge dynamics observed in product quality management have been reported to reduce costs more effectively (Cusumoto et al., 1995; Kono & Nonaka, 1995; Madhavan & Grover, 1998). Accordingly, it is possible that companies that require a simultaneous involvement of product leaders and functional staff in target cost allocation decisions may experience, and thus expect, a higher level of cost reduction performance. We formulate this reasoning as the third hypothesis and state:

H3. Companies that require a simultaneous involvement of product leaders and functional staff in target cost allocation decisions would show a higher level of cost reduction performance.

THE METHODOLOGY

The Survey

The Tokyo Stock Exchange lists 518 companies in the four industries (machinery, electronics equipment, transportation equipment, and precision instruments) identified in target cost management literature as the Japanese manufacturers using target cost management (Tani et al., 1994; Nishizawa, 1995). A questionnaire survey of the 518 firms was performed regarding

their cost management practices. Forty-three questions were asked in the areas of: (1) the attributes of the firm and target cost management; (2) target cost management organizations; (3) methods for setting target sales price, target profit, and target cost; target allocation, target achievement level, and target cost follow-up: (4) value engineering operations in product design; (5) utilization of cost tables; (6) product preparation and supplier relationship; and (7) uncertainties and complexities in the external environment.

The four specific questions on the target cost management practices were: (1) the organizational relationship between product development and product design; (2) the decision authority on target cost allocation; (3) the method used to set target cost; and (4) the level of cost reduction performance expected. The target cost management part of the survey questionnaire is presented in Appendix A. The overall response rate was 28%. The response rates by industry are shown in Table 1. In this survey, three target cost-setting methods were used: the subtraction method, the hybrid method, and the addition method. The three methods and the levels of cost reduction performance expected by the firm were identified based on the extensive pre-survey interviews conducted with cost accounting practitioners in the industries represented. The methods and the levels were generally agreed to be the norms by the participating practitioners.

ANALYSIS OF THE RESULTS

Table 2 presents the descriptive statistics for the variables. Table 3 shows the alpha reliability coefficients of this survey. The Cronbach's α of the "target cost organizations (section 2)" and "target cost details (section 3)" are above 0.58, which means the convergence of the variables should be considered reasonable. In addition, the factor loadings on the variables are

Table 1. Response Rates by Industry.

Type	Mailed	Received	Response Rate (%)	Adjusted Response Rate (%)
Machinery	194	49	25.26	24.74
Electronics	198	49	24.74	24.24
Transportation	90	38	42.22	41.11
Precision	36	10	27.78	27.78
Total	518	146	28.19	27.61

Table 2. Descriptive Statistics of Research Variables.

Variables	*n*	Mean	SD	Range
Product development org. (Q10)	136	1.779	0.795	1–3
Authority to allocate target cost (Q11)	137	2.803	0.976	1–5
Cost-setting method (Q18)	130	2.131	0.751	1–3
Cost-reduction performance (Q20)	134	3.440	1.051	1–5

Table 3. Reliability Analysis.

Sections	Cronbach's α
(1) Attributes of the company and implantation of target costing	0.457
(2) Target costing organizations	0.909
(3) Methods for setting target sales prices, sales profits, and target costs; allocation, achievement level, and follow-up of target costs.	0.588
(4) VE operations in product design	0.934
(5) Utilization of cost tables	0.893
(6) Product preparation and suppliers' relationship	0.913
(7) Uncertainties and complexities of the external environment	0.448

above 0.5, supporting the reasonableness of the construct validity of this study. A four-way contingency table for the four factors contained in the four questions is presented in Table 4. The log-linear models were constructed based on the four variables *Org*, *All*, *Set*, and *Per*, and the categories corresponding to the respective variables i, j, k, and l ($i = 1, 2$; $j = 1, 2, 3$; $k = 1, 2, 3$; $l = 1, 2, 3$) are presented in Table 5. For the 54 cells in the responses, the expected frequency for each cell was f_{ijkl}, and the observed frequency was g_{ijkl}.

The log-linear model includes the intercept, all main-effect terms (λ_i^{Org}, λ_k^{All}, λ_j^{Set}, and λ_l^{Per}), and interaction terms ($\lambda_{ik}^{Org\cdot Set}$, $\lambda_{ij}^{Org\cdot All}$, $\lambda_{il}^{Org\cdot Per}$,...) as shown below:

$$\begin{aligned}\log f_{ijkl} = {} & \lambda + \lambda_i^{Org} + \lambda_j^{All} + \lambda_k^{Set} + \lambda_l^{Per} + \lambda_{ij}^{Org\cdot All} + \lambda_{ik}^{Org\cdot Set} + \lambda_{il}^{Org\cdot Per} \\ & + \lambda_{jk}^{All\cdot Set} + \lambda_{jl}^{All\cdot Per} + \lambda_{kl}^{Set\cdot Per} + \lambda_{ijk}^{Org\cdot All\cdot Set} + \lambda_{ijl}^{Org\cdot All\cdot Per} \\ & + \lambda_{ikl}^{Org\cdot Set\cdot Per} + \lambda_{jkl}^{All\cdot Set\cdot Per} + \lambda_{ijkl}^{Org\cdot All\cdot Set\cdot Per} \end{aligned} \tag{6}$$

$$(i = 1, \ldots, 2;\ j = 1, \ldots, 3;\ k = 1, \ldots, 3;\ l = 1, \ldots, 3)$$

Table 4. Factor Loading of Research Variables.

Variables	Factor Loading
Product development structure (Q10)	0.645
Authority to allocate target cost (Q11)	0.590
Cost-setting method (Q18)	0.516
Cost-reduction performance (Q20)	0.598

Table 5. Categories of Factors (*Org*, *All*, *Set*, and *Per*).

Factor	Item	Category	Content
Org	Product development organization (Q10)	$i = 1$	Matrix organization (Answer 1)
		$i = 2$	Multi-center organization (Answer 2)
All	Authority to allocate target cost (Q11)	$j = 1$	Primarily by project leaders (Answer 1)
		$j = 2$	Simultaneously by project leaders and functional staff (Answer 2)
		$j = 3$	Primarily by functional staff (Answer 3)
Set	Method used to set target cost (Q18)	$k = 1$	Subtraction method (Answer 1)
		$k = 2$	Hybrid method (Answer 2)
		$k = 3$	Addition method (Answer 3)
Per	Cost reduction performance (Q20)	$l = 1$	Below 70% (Answers 1 and 2)
		$l = 2$	71~80% level (Answer 3)
		$l = 3$	Above 81% (Answers 4 and 5)

The other parameters are defined in a similar fashion. All parameters estimated are shown in Table 6. For the simplicity of the model, we obtain unsaturated models by considering the parameters that show the asymptotic t-values of over ± 1 (Appendix B). In the unsaturated models, certain parameters are set to zero, as in Eq. (6). Certain higher-order parameters are treated as zero, according to the hierarchical principles. For instance, if $\lambda_{kl}^{Set \cdot Per} = 0$, then

$$\lambda_{ikl}^{Org \cdot Set \cdot Per} = \lambda_{jkl}^{All \cdot Set \cdot Per} = \lambda_{ijkl}^{Org \cdot All \cdot Set \cdot Per} = 0$$

Therefore, the log-linear model in the case is stated as follows:

$$\begin{aligned} \log\ f_{ijkl} = \lambda + \lambda_i^{Org} + \lambda_j^{All} + \lambda_k^{Set} + \lambda_l^{Per} + \lambda_{ij}^{Org \cdot All} + \lambda_{ik}^{Org \cdot Set} + \lambda_{il}^{Org \cdot Per} \\ + \lambda_{jk}^{All \cdot Set} + \lambda_{kl}^{All \cdot Per} + \lambda_{ijk}^{Org \cdot All \cdot Set} + \lambda_{ijl}^{Org \cdot All \cdot Per} \end{aligned} \quad (7)$$

Table 6. Saturated Model for Factors (*Org*, *All*, *Set*, and *Per*).

Parameter	Coefficient	Asymptotic *t*-value	Parameter	Coefficient	Asymptotic *t*-value
λ_2^{Org}	−0.035	−0.141	$\lambda_{222}^{Org\cdot All\cdot Set}$	**0.788**	**1.000**
λ_2^{All}	−0.152	−0.552	$\lambda_{223}^{Org\cdot All\cdot Set}$	0.449	0.571
λ_3^{All}	**−0.645**	**−2.029**	$\lambda_{232}^{Org\cdot All\cdot Set}$	0.281	0.327
λ_2^{All}	**0.668**	**2.120**	$\lambda_{233}^{Org\cdot All\cdot Set}$	0.159	0.202
λ_3^{Set}	**0.415**	**1.273**	$\lambda_{222}^{Org\cdot All\cdot Per}$	−0.791	−0.589
λ_2^{Per}	**0.498**	**1.493**	$\lambda_{223}^{Org\cdot All\cdot Per}$	**−1.515**	**−1.051**
λ_3^{Per}	**0.881**	**2.835**	$\lambda_{232}^{Org\cdot All\cdot Per}$	−1.072	−0.639
$\lambda_{22}^{Org.All}$	−0.036	−0.065	$\lambda_{233}^{Org\cdot All\cdot Per}$	−0.050	−0.030
$\lambda_{23}^{Org.All}$	0.274	0.430	$\lambda_{222}^{All\cdot Set\cdot Per}$	**2.650**	**1.750**
$\lambda_{22}^{Org.Set}$	0.002	0.002	$\lambda_{223}^{All\cdot Set\cdot Per}$	**1.559**	**1.212**
$\lambda_{23}^{Org.Set}$	−0.641	−0.984	$\lambda_{232}^{All\cdot Set\cdot Per}$	**2.396**	**1.429**
$\lambda_{22}^{Org.Per}$	−0.601	−0.902	$\lambda_{233}^{All\cdot Set\cdot Per}$	0.547	0.332
$\lambda_{23}^{Org.Per}$	−0.577	−0.929	$\lambda_{322}^{All\cdot Set\cdot Per}$	−0.055	−0.033
$\lambda_{22}^{All.Set}$	−0.054	−0.081	$\lambda_{323}^{All\cdot Set\cdot Per}$	−0.450	−0.286
$\lambda_{23}^{All.Set}$	**−0.989**	**−1.374**	$\lambda_{332}^{All\cdot Set\cdot Per}$	−0.903	−0.525
$\lambda_{32}^{All.Set}$	0.536	0.639	$\lambda_{333}^{All\cdot Set\cdot Per}$	−0.435	−0.276
$\lambda_{33}^{All.Set}$	0.109	0.133	$\lambda_{2222}^{Org.All.Set.Per}$	−0.213	−0.114
$\lambda_{22}^{All.Per}$	−0.676	−0.893	$\lambda_{2223}^{Org.All.Set.Per}$	−0.718	−0.461
$\lambda_{23}^{All.Per}$	−0.329	−0.512	$\lambda_{2232}^{Org.All.Set.Pet}$	1.103	0.553
$\lambda_{32}^{All.Per}$	0.353	0.421	$\lambda_{2233}^{Org.All.Set.Per}$	−0.010	−0.005
$\lambda_{33}^{All.Per}$	−0.131	−0.160	$\lambda_{2322}^{Org.All.Set.Per}$	0.553	0.251
$\lambda_{22}^{Set.Per}$	−0.035	−0.141	$\lambda_{2323}^{Org.All.Set.Per}$	1.822	0.834
$\lambda_{23}^{Set.Per}$	−0.152	−0.552	$\lambda_{2332}^{Org.All.Set.Per}$	0.130	0.061
$\lambda_{32}^{Set.Per}$	**−0.645**	**−2.029**	$\lambda_{2333}^{Org.All.Set.Per}$	0.963	0.471
$\lambda_{33}^{Set.Per}$	**0.668**	**2.120**			

This is a hierarchical model. For simplicity, this model is assigned a symbol that lists the highest-order effects as *Org-All-Set*/*Org-All-Per*. This symbol is the definition set of the log-linear model of formula (2). For precise estimations of parameters and cell probabilities in the log-linear model, model fitting is performed. We use partial chi-square statistics to test that the expected frequencies f_{ijkl} for a given model fit the observed frequencies g_{ijkl}. We also employ likelihood ratio statistics and the Akaike Information Criterion (*AIC*) to measure the model goodness of fit.

Under the parsimony principle, we choose the model with the smallest number of free parameters when there are several models with similar

maximum likelihood values. To facilitate the comparison of the goodness of fit between the saturated and hierarchical models, we compute the difference between the two models as follows:

$$AIC' = AIC_H - AIC_S = G^2 - 2df \tag{8}$$

where AIC_H is the AIC of the hierarchical model, AIC_S the AIC of the saturated model, and df the degrees of freedom.

Similarly, we choose the model with the smallest AIC'. The likelihood ratio statistic of the selected model $\{Org\text{-}All\text{-}Per/All\text{-}Set\}$ is statistically significant at the 0.875 level and the AIC' value of this model is the smallest (Table 7). Accordingly, the hierarchical model is formulated as follows:

$$\begin{aligned}\log\ f_{ijkl} = \lambda + \lambda_i^{Org} + \lambda_j^{All} + \lambda_k^{Set} + \lambda_l^{Per} + \lambda_{ij}^{Org\cdot All} + \lambda_{il}^{Org\cdot Per} \\ + \lambda_{jk}^{All\cdot Set} + \lambda_{jl}^{All\cdot Per} + \lambda_{ijl}^{Org\cdot All\cdot Per}\end{aligned} \tag{9}$$

We set the expected frequency f_{1131} for cell (Org_1, All_1, Set_3, Per_1) to be a criterion to estimate the other parameters. The estimated parameters are presented in Table 8. According to the alternative constraints, the parameters satisfy the following constraints:

$$\lambda_1^{Org} = \lambda_1^{All} = \lambda_3^{Set} = \lambda_1^{Per} = 0$$

$$\lambda_{1j}^{Org\cdot All} = \lambda_{i1}^{Org\cdot All} = \lambda_{1l}^{Org\cdot Per} = \lambda_{i1}^{Org\cdot Per} = \lambda_{1l}^{All\cdot Per} = \lambda_{j1}^{All\cdot Per} = \lambda_{1k}^{All\cdot Set} = \lambda_{j3}^{All\cdot Set} = 0$$

$$\lambda_{1jl}^{Org\cdot All\cdot Per} = \lambda_{i1l}^{Org\cdot All\cdot Per} = \lambda_{ij1}^{Org\cdot All\cdot Per} = 0$$

$$i = 1, 2;\ j = 1, \ldots, 3;\ k = 1, \ldots, 3;\ l = 1, \ldots, 3$$

Table 7. Goodness-of-Fit Tests for Loglinear Models (*Org*, *All*, *Set*, and *Per*).

Model	Set of the Hierarchical Model	df	G^2 (p-value)	AIC′
1	*Org-All-Set/Org-All-Per/All-Set-Per*	12	5.882 (0.922)	−18.118
2	*Org-All-Per/All-Set-Per/Org-Set*	16	9.223 (0.904)	−22.774
3	*Org-All-Set/All-Set-Per*	18	16.611 (0.550)	−19.389
4	*Org-All-Per/Org-Set/All-Set/Set-Per*	24	17.822 (0.812)	−30.180
5	*Org-All-Set/Org-All-Per*	24	14.579 (0.932)	−33.421
6	*All-Set-Per/Org*	26	24.010 (0.575)	−27.990
7	*Org-All-Per/Org-Set/All-Set*	28	18.933 (0.900)	−37.067
8	***Org-All-Per/All-Set***	**30**	**21.411 (0.875)**	**−38.589**
9	*Org-All-Per/Set-Per*	30	28.980 (0.519)	−31.102
10	*Org-All-Set/Set-Per*	30	27.303 (0.607)	−32.697

Table 8. Parameter Estimation for Model (*Org-All-Per/All-Set*).

$$\log f_{ijkl} : \log f_{1131} = \lambda + \lambda_i^{Org} + \lambda_j^{All} + \lambda_k^{Set} + \lambda_l^{Per} + \lambda_{ij}^{Org\cdot All} + \lambda_{il}^{Org\cdot Per} + \lambda_{jk}^{All\cdot Set} + \lambda_{jl}^{All\cdot Per} + \lambda_{ijl}^{Org\cdot All\cdot Per} \quad i = 2;\ j = 2,3;\ k = 1,2;\ l = 2,3$$

λ_i^{Org}			$i = 2$	
			−0.830	
			(−0.817)	
λ_k^{All}			$j = 2$	$j = 3$
			−0.729	−7.649
			(−0.617)	(−0.315)
λ_j^{Set}			$k = 1$	$k = 2$
			−0.754	−0.194
			(−1.758)	(−0.538)
λ_l^{Per}			$l = 2$	$l = 3$
			2.398**	**2.398****
			(2.296)	**(2.296)**
$\lambda_{ij}^{Org\cdot All}$			$j = 2$	$j = 3$
		$i = 2$	−1.897	5.771
			(−1.421)	(0.237)
$\lambda_{il}^{Org\cdot Per}$		$i = 2$	$l = 2$	$l = 3$
			−**2.909****	−1.928*
			(−**2.282**)	(−1.620)
$\lambda_{jl}^{All\cdot Per}$		$j = 2$	$l = 2$	$l = 3$
			−**2.398****	−1.705
			(−**1.901**)	(−1.408)
		$j = 3$	6.081	6.774
			(0.250)	(0.279)
$\lambda_{jk}^{All\cdot Set}$			$k = 1$	$k = 2$
		$j = 2$	1.110*	0.887†
			(1.700)	(1.512)
		$j = 3$	−1.038	0.800
			(−0.893)	(1.285)
$\lambda_{ijl}^{Org\cdot All\cdot Per}$			$l = 2$	$l = 2$
	$i = 2$	$j = 2$	**3.196****	2.216†
		$j = 3$	**(1.943)**	(1.477)
			−3.961	−6.146
			(−0.163)	(−0.250)

$^{**}p<0.01$, $^{*}p<0.05$, $^{\dagger}p<0.1$.

Note: Values in parentheses () denote asymptotic *t*-value.

The empirical results indicate a distinct relationship pattern. As Table 4 indicates, companies with a multi-center organization of product development (*Org*) do not particularly prefer the use of any particular target cost-setting method (*Set*) in the tested model {*Org-All-Per*/*All-Set*}. However, the parameters of $\{\lambda_{21}^{All\cdot Set}\}$ and $\{\lambda_{22}^{All\cdot Set}\}$ are significantly positive ($\lambda_{21}^{All\cdot Set} = 1.110$ (asymptotic t-value $= 1.700$); $\lambda_{22}^{All\cdot Set} = 0.887$ (asymptotic t-value $= 1.512$)). This means that companies with a market orientation strategy prefer a close alignment of target cost allocation authority with a market-oriented target cost-setting method. H2 is supported and H1 is not.

We find a relationship between the use of a subtraction method, which is the most market-oriented among the three methods, and the simultaneous involvement of product managers and functional staff in target cost allocation decisions. The results indicate that the use of a subtraction method is correlated with the simultaneous involvement of product managers and functional staff in target cost allocation decisions, but is not correlated with the choice of a product development organization.

Table 8 also shows that the parameters of $\{\lambda_{222}^{Org\cdot All\cdot Per}\}$ and $\{\lambda_{223}^{Org\cdot All\cdot Per}\}$ are positive ($\lambda_{222}^{Org\cdot All\cdot Per} = 3.196$ (asymptotic t-value $= 1.943$); $\lambda_{223}^{Org\cdot All\cdot Per} = 2.216$ (asymptotic t-value $= 1.477$)). When project leaders and functional staff get simultaneously involved in target cost allocation decisions, higher levels of cost reduction performance are expected in a multi-centered organization of product development. Companies with higher levels of cost reduction performance tend to make project leaders and functional staff get simultaneously involved in target cost allocation decisions. Perhaps cost reduction activities are facilitated in such a setting, because the leaders in each center are responsible for maintaining personnel relationships within the organization and motivating their staff to achieve the organizational goals (House, 1971; Steers, 1975; Isabella & Waddock, 1994; Durham, Knight, & Locke, 1997).

SUMMARY AND CONCLUSIONS

In the current target cost management study, we find that organizational efficiency is connected with cost reduction performance from the two perspectives: (1) the relationship between the simultaneous involvement of project leaders and functional staff and the use of a particular target cost-setting method (subtraction method), and (2) the correlation of the simultaneous involvement of project leaders and functional staff with the level of expected cost reduction performance. There is no connection

between product development organization and the choice of a particular target cost-setting method. There is, however, a connection between target cost allocation authority (simultaneous involvement of project leaders and functional staff) and the use of a target cost-setting method (subtraction method).

The simultaneous involvement of project leaders and functional staff in target cost allocation decisions contributes to higher levels of cost reduction performance. Companies with higher levels of cost reduction performance (expected or experienced) tend to make project leaders and functional staff get simultaneously involved in target cost allocation decisions.

IMPLICATIONS FOR FUTURE RESEARCH

Target cost management research has advanced to a level where a study can determine if and how organizational learning is affected by the implementation of target cost management. This status indicates that target cost management is not an isolated management control process. It is related to organizational alignment (in connection with competitive strategy), factors that drive operations management's (supply chain management included) use of management accounting information, comprehensive benchmarking system, and inter-organizational cost management systems.

A recent survey (Pierce, 2002) reveals that about 40% of responding organizations in the U.S. has adopted target cost management. The findings in the current study, based on Japanese manufacturers' responses, could be tested with the U.S. counterparts as research subjects. Environmental or cultural differences in the responses, if found, would be relevant subjects for future research. The findings in our study may signal important differences involving product development organization and the way product development decision authority is formulated. The fact that our first (second) hypothesis was not (was) supported may imply such possibility. With respect to the issues of product development structure, management accounting researchers need more work on how the organizational learning that deals with such structure is different from the single- and the double-loop learning theorized by Argyris (1977).

Based on our findings, it is conceivable that companies using a multi-center organization can operate with a different mode of target cost allocation decision process. It would be fruitful to examine how a multi-center organization can support an effective organizational learning within the organizational learning theory (Romme, 1996). Future research would

shed more light on whether it is the simultaneous involvement of functional staff in allocating target costs or the organizational learning of functional staff involving current production technology, production equipment operation, related costs, and the lead time required to develop new products that contributes more to the attainment of target costs.

NOTES

1. A Japanese target costing survey (Lee & Monden, 1996) reported that about 45% of the responding companies used the matrix organization of product development and 32% used the multi-center product development organization.

REFERENCES

Ansari, S. L., Bell, J. E., & the Target Cost Core Group. (1997). *Target costing.* New York: McGraw-Hill.

Argyris, C. (1977). Double loop learning in organizations. *Harvard Business Review*(September–October), 59–72 AQ: 5.

Choe, J. (2002). The organizational learning effects of management accounting information under advanced manufacturing technology. *European Journal of Information Systems*, *11*, 142–158.

Clark, K. B., & Fujimoto, T. (1991). *Product development performance*. Boston: Harvard Business School Press.

Cooper, R., & Slagmulder, R. (1997). *Target costing and value engineering*. Portland, OR: Productivity.

Cusumano, M. A., & Nobeoka, K. (1998). *Thinking beyond lean: How multi-project management is transforming product development at Toyota and other companies*. New York: Free Press.

Cusumoto, K., Nonaka, Y., & Nagata, K. (1995). Japanese companies' organizational capability in new product development (in Japanese). *Organizational Science (Soshiki Kakaku)*, *29*, 92–108.

Davis, S. M., & Lawrence, P. R. (1977). *Matrix*. Reading, MA: Addison-Wesley.

Durham, C. C., Knight, D., & Locke, E. A. (1997). Effects of leader role, team-set goal difficulty, efficacy, and tactics on team effectiveness. *Organizational Behavior and Human Decision Processes*, *72*, 203–231.

Euske, K. J., Lebas, M. J., & McNair, C. J. (1993). Performance measurement in an international setting. *Management Accounting Research*, *4*, 275–300.

House, R. J. (1971). A path-goal theory of leader effectiveness. *Administrative Science Quarterly*, *16*, 321–338.

Isabella, L. A., & Waddock, S. A. (1994). Top management team certainty: Environmental assessments, teamwork, and performance implications. *Journal of Management*, *20*, 835–858.

Kanazawa, Y., & Monden, Y. (1994). Target sales pricing for target costing system (in Japanese). *The Journal of Management Accounting, Japan*, *3*, 3–25.
Kloot, L. (1997). Organizational learning and management control systems: Responding to environmental change. *Management Accounting Research*, *8*, 47–73.
Kondo, K. (1990). The transfiguration of cost management: From technology-oriented paradigm to market-oriented paradigm (in Japanese). *Kaikei (Accounting)*, *137*, 64–78.
Kono, N., & Nonaka, Y. (1995). *Chiryoku Keiei (Intellectual Business)*. Nihon Keizai: Shinbun Sha.
Lee, C. H., & Monden, Y. (1996). *Target costing survey of Japanese manufacturers*. Tsukuba, Japan: Institute of Socio-Economic Planning, University of Tsukuba.
Madhavan, R., & Grover, R. (1998). From embedded knowledge to embodied knowledge: New product development as knowledge management. *Journal of Marketing*, *62*(October), 1–12.
Monden, Y., & Hamada, K. (1991). Target costing and kaizen costing in Japanese automobile companies. *Journal of Management Accounting Research*, *3*, 16–34.
Nishizawa, O. (1995). *Nihon Kigyo No Kanri Kaikei – Shuyo 229 Sha No Jittai Bunseki* [Managerial accounting in Japanese companies: Empirical analysis of 229 companies], Tokyo: Tsuo Keizai Sha.
Pierce, B. (2002). Target cost management: comprehensive benchmarking for a competitive market. *Accountancy (Ireland)*, *April*, 30–32.
Powell, T. C. (1992). Organizational alignment as competitive advantage. *Strategic Management Journal*, *13*, 119–134.
Romme, G. (1996). Making organizational learning work: Consent and double linking between circles. *European Management Journal*, *14*(1), 69–75.
Shibata, S., & Nakahashi, K. (1997). *The theory and practice of business administration*. Tokyo: Keizei Jyoho Shubbansya (in Japanese).
Steers, R. M. (1975). Task-goal attributes, *n* achievement, and supervisory performance. *Organizational Behavior and Human Performance*, *13*, 392–403.
Tanaka, M. (1989). Cost planning and control systems in the design phase of a new product. In: Y. Monden & M. Sakurai (Eds), *Japanese management accounting* (pp. 49–72). Portland, OR: Productivity.
Tani, T., Okano, H., Shimizu, N., Iwabuchi, Y., Fukuda, J., & Cooray, S. (1994). Target costing management in Japanese companies: Current state of the art. *Management Accounting Research*, *5*(67), 81.
Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method. *Biometrika*, *61*(439), 447.

APPENDIX A:. SURVEY QUESTIONNAIRE – TARGET COST MANAGEMENT PORTION

Q10 What was the organizational relationship between product development and product design when they were established in your firm?

1. Product development and product design were established as independent and separate departments.

2. Product design was established in product development department based on the functional parts designed.
3. Others.

Q11 Who makes the decision on how to allocate the target cost of functional parts or components?

1. Primarily product managers.
2. Simultaneously by product managers and design staff.
3. Primarily design staff.

Q18 What method is used to set target cost in your firm?

1. Subtraction method: Target cost is set as target profit is subtracted from target sales price.
2. Hybrid method: Target cost is set as the "allowable cost" is compared to the "estimated cost."
3. Addition method: Target cost is set as the added-up total of the unit "estimated cost" minus the per-unit profit improvement target.

Q20 What level of cost reduction performance was achieved for the target cost of the product recently developed?

1. 60% (or below).
2. 61 to $\sim$ 70%
3. 71 to $\sim$ 80%
4. 81 to $\sim$ 90%
5. 91 to $\sim$ 100% (or above).

APPENDIX B. LOG-LINEAR MODELS

The log-linear models were constructed based on the four variables of *Org*, *All*, *Set*, and *Per* in the current study. The expected frequency for each cell was f_{ijkl}, and the observed frequency was g_{ijkl} for each cell in the cross-table. The log-linear model is similar to the models of a four-way factorial analysis of variance (ANOVA). The first parameters are defined as follows:

$$\mu = \log f_{i^*j^*k^*l^*}$$

$$\lambda_i^{Org} = \log f_{ij^*k^*l^*} - \mu, \quad \lambda_j^{All} = \log f_{i^*jk^*l^*} - \mu$$
$$\lambda_k^{Set} = \log f_{i^*j^*kl^*} - \mu, \quad \lambda_l^{Per} = \log f_{i^*j^*k^*l} - \mu$$

$$\lambda_{ijk}^{Org\cdot All\cdot Set} = \log f_{ijkl^*} - (\lambda_{ij}^{Org\cdot All} + \lambda_{ik}^{Org\cdot Set} + \lambda_{il}^{Org\cdot Per}) - \mu$$
$$\lambda_{ijl}^{Org\cdot All\cdot Per} = \log f_{ijk^*l} - (\lambda_{ij}^{Org\cdot All} + \lambda_{il}^{Org\cdot Per} + \lambda_{jl}^{All\cdot Per}) - \mu$$
$$\lambda_{ikl}^{Org\cdot Set\cdot Per} = \log f_{ij^*kl} - (\lambda_{ik}^{Org\cdot Set} + \lambda_{il}^{Org\cdot Per} + \lambda_{kl}^{Set\cdot Per}) - \mu$$
$$\lambda_{jkl}^{All\cdot Set\cdot Per} = \log f_{i^*jkl} - (\lambda_{jk}^{All\cdot Set} + \lambda_{jl}^{All\cdot Per} + \lambda_{kl}^{Set\cdot Per}) - \mu$$

$$\begin{aligned}\lambda_{ijkl}^{Org\cdot All\cdot Set\cdot Per} = {} & \log f_{ijkl} - (\lambda_{ijk}^{Org\cdot All\cdot Set} + \lambda_{ijl}^{Org\cdot All\cdot Per} + \lambda_{ikl}^{Org\cdot Set\cdot Per} + \lambda_{jkl}^{All\cdot Set\cdot Per}) \\ & - (\lambda_{ij}^{Org\cdot All} + \lambda_{ik}^{Org\cdot Set} + \lambda_{il}^{Org\cdot Per} + \lambda_{jk}^{All\cdot Set} + \lambda_{jl}^{All\cdot Per} + \lambda_{kl}^{Set\cdot Per}) \\ & - (\lambda_{i}^{Org} + \lambda_{j}^{All} + \lambda_{k}^{Set} + \lambda_{l}^{Per}) - \mu\end{aligned}$$

According to Wedderburn (1974), the quasi-likelihood estimation of the dispersion parameter ϕ is

$$\phi = X^2/(N - df)$$

where

$$X^2 = \sum_2 \sum_3 \sum_3 \sum_3 (g_{ijkl} - f_{ijkl})^2 / f_{ijkl}$$

N is the number of cells, and df the degrees of freedom.

We assure the efficiency of the coefficients between variance and mean by using the dispersion parameter ϕ, so that asymptotic t-value = coefficient/$\sqrt{\phi}$

Some of the hierarchical (or unsaturated) models are structured based on the parameters that show the asymptotic t-values of over ± 1. For goodness of fit, one hierarchical model that has the lowest value of *AIC* should be selected for parameter estimation.

One expected frequency $f_{i^*j^*k^*l^*}$ is set to be the criterion for the estimation of the other parameters. For parameter estimation, the iterative proportional fitting (IPF) algorithm is used to solve likelihood equations. The IPF algorithm for maximum likelihood estimation generates a sequence of fitted values converging to a unique solution that satisfies the model and matches the sufficient statistics.

This page intentionally left blank

ANTECEDENTS AND CONSEQUENCES OF BUDGET PARTICIPATION

Adam S. Maiga

ABSTRACT

This research uses structural .equation modeling to investigate the relationships between environmental uncertainty, budget communication, budget influence, budget goal commitment and managerial performance. To this end, data from 173 U.S. individual managers were used for the study. The results show that environmental uncertainty significantly affects both budget communication and budget influence which, in turn, impact budget goal commitment. Also, budget goal commitment is significantly related to managerial performance. Implications, limitations and directions for future research are discussed.

INTRODUCTION

The importance of budget participation as a means of improving performance has been studied extensively in the accounting literature (Brownell, 1981, 1982; Mia, 1989; Kren, 1992; Magner, Welker, & Campbell, 1996; Nouri & Parker, 1996). It is suggested that budget participation also serves an informational function, whereby subordinates can gather, exchange and

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 211–231

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14010-6

disseminate job-relevant information to facilitate their decision-making process and to communicate their private information to organizational decision-makers (Murray, 1990; Kren, 1992; Shields & Young, 1993; Magner et al., 1996; Nouri & Parker, 1996). Furthermore, it is argued that the act of participation in the budgeting process serves a function by inducing subordinates to accept and commit to their budget goals (Hofstede, 1968; Kenis, 1979; Merchant, 1981).

However, empirical findings regarding the direct association between budget participation and performance have proved to be mixed, ranging from strongly positive (Brownell & McInnes, 1986) to weak (Milani, 1975) to nonexistent (Kenis, 1979) and even negative (Mia, 1988; Brownell, 1981). Greenberg, Greenberg, and Nouri's (1994) meta-analytical results indicate that differences in research methods are not responsible for these diverse findings. The literature suggests that the relation between participation and performance is influenced by moderating variables as well as mediating one (Shields & Shields, 1998). In addition, the main focus of budget participation research since Brownell's (1982) review has still been on its consequences rather than its antecedents (Binberg, Shields, & Young, 1990). Shields and Young (1993) propose that in order for research to make progress toward increasing an understanding of participative budgeting, it should first focus on explaining participative budgeting's antecedents. Furthermore, almost all prior studies on budgetary processes have been based on Milani's (1975) measure of budget participation and they have viewed budgetary participation as a unidimensional construct even though Milani (1975) suggested the possibility of two dimensions, and Brownell (1982) and Hassel and Cunningham (1993, 1996) isolated two dimensions: communication and influence.

This paper contributes to the budget participation–performance literature in the following ways. First, this study continues the example of Hassel and Cunningham (1993, 1996) and examines budget communication and influence separately. Because the two dimensions are conceptually independent, it is appropriate to develop hypotheses separately for each dimension. Factor analysis, presented below, tends to confirm the existence of these two dimensions. Second, I propose an integrated research model (see Fig. 1) that incorporates and examines the relationships among environmental uncertainty, budget participation (i.e., budget communication and budget influence),[1] budget goal commitment and managerial performance. Specifically, I hypothesize that environmental uncertainty increases both budget communication and budget influence. Budget communication and budget influence, in turn, trigger subordinates' commitment to budget goal. The degree of budget goal commitment is expected to increase managerial

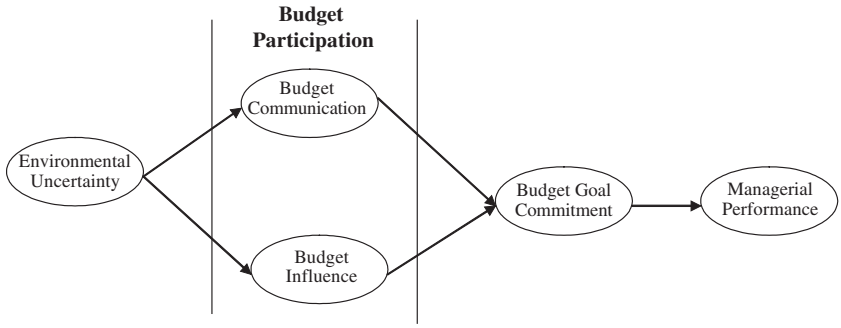


Fig. 1. Research Model.

performance. It is expected that an examination of a structural model that incorporates these variables will improve our understanding of the budget participation–performance linkage. The structural equation modeling technique allows us (1) to simultaneously test the linkages in the budget participation–performance structural model and (2) the specification of more complex models and takes into account measurement errors.

The remainder of the paper is structured as follows. In the next section, definitions, literature review and hypotheses are developed. This is followed by the research method and results of the hypothesis testing. Finally, directions for future research are discussed.

DEFINITIONS, LITERATURE REVIEW AND HYPOTHESES

The objective of this research is to examine the effects of environmental uncertainty on budget participation (budget communication and budget influence), the impact of budget participation on budget goal commitment and the effect of budget goal commitment on managerial performance. In this section, these variables are first defined, followed by the literature review and identification of hypotheses to be tested. The overall conceptual framework is illustrated in Fig. 1.

Definitions

This section defines key terms in this study, reviews the literature and derives the hypotheses to be tested.

One factor that influences budgetary decisions is the decision-maker's uncertainty resulting from the complexity of the environment. Uncertainty is defined as perceived uncertainty with respect to environmental factors such as customers, suppliers, competitors, government and technology (Ibrahim, 1993; Govindarajan, 1986). The cause of this uncertainty for decision-makers is the complexity of the environment. The complexity of the environment can change, thereby increasing or decreasing the perceived uncertainty (Tichy, 1982).

Budget participation is defined as a means of communication and influence of managers in the budgetary process and the extent of their influence over the setting of budgetary targets (Milani, 1975; Brownell, 1982; Hassel & Cunningham, 1993, 1996). Accordingly, this study indicates that budgetary is at least bi-dimensional: (1) the extent to which information is communicated back and forth between headquarters and subunit managers and (2) the extent to which subunit managers influence the budgets for their inputs. For example, subunits communication with headquarters may be extensive but the budget may be set by headquarters with no influence by the subordinate managers. By contrast, in highly autonomous subunits, subordinates may have substantial influence over their budgets with only limited communication with headquarters (Hassel & Cunningham, 1996).

Budget communication is the extent to which information is exchanged between superiors and subunit managers about factors that affect the budget. *Budget influence* is the degree to which subunit managers perceive that they have command over the process that establishes the criteria under which they may be evaluated (Hassel & Cunningham, 1993).

Budget goal commitment is defined in this study as the determination to try for a budget goal and the persistence in pursuing it over time (Locke, Shaw, Saari, & Latham, 1981). It is characterized by first, a strong belief in and acceptance of the budget goals, and a willingness to exert considerable effort on behalf of the organization.

Managerial performance refers to the degree of successful achievement attained by an employee (Ferris, 1977; Mahoney, Jerdee, & Carroll, 1963).

Literature Review and Hypotheses Development

Environmental Uncertainty and Budget Participation

Agency theorists argue that the demand for participative budgeting arises because various parties engaged in the budgeting process possess differential

information about uncertainty (e.g., central and local management) (Christensen, 1982; Baiman & Evans, 1983; Penno, 1990; Kerby, Reichelstein, Sen, & Piak, 1991). Participation provides a means of pooling the experience and knowledge of budgeted managers and their superiors and offers the potential to assist in resolving these uncertainties. Lawrence and Lorsch (1967) found that firms facing high environmental uncertainty tended to allow greater participation. Govindarajan (1986) and Hopwood (1976) extended this reasoning about participation in decision-making in general to participation in budgeting and suggested that greater budgetary participation should be found in organizations facing greater environmental uncertainties. This argument is supported by Simons (1987) who argues that when environmental conditions are unstable, communication around the budget has to be increased. Empirical tests tend to support these arguments (Brownell, 1985; Mia, 1989). Therefore, communication in the budget process serves a major information exchange role about critical factors of the environment. As a result, it can be argued that as environmental uncertainty increases, budget communication increases. More specifically,

H1. Environmental uncertainty is positively related to budget communication.

Becker and Gordon (1966) suggest that where interaction between an organization and its environment is complex and rapid specification of procedures is required within the organization, management is forced to allow lower hierarchic levels to specify their own procedures. This suggests that when the environment is dynamic, managers must have some degree of autonomy over the budget process. Hence, the link between environmental uncertainty and budget influence is expected to become stronger when environmental uncertainty is high because (1) managers are the ones most familiar with their budget environment and (2) budgets imposed by the headquarters would not reflect subunit conditions. Thus, it may follow that managers operating in uncertain situations will have more autonomy over decision-making relevant to their subunits. As a result, under environmental uncertainty, managers of local subunits must be able to have influence over the budget participation process.

The above argument can be summarized as follows:

H2. Environmental uncertainty is positively related to budget influence.

Budget Participation and Budget Goal Commitment

Recent a priori arguments suggest that benefits from budget participation are derived primarily from the communication of information (Miller & Monge,

1986; Vroom & Jago, 1988). Similarly, in their concluding discussion, Brownell and Dunk (1991) suggest that budgetary participation serves as a major information exchanging role and that the primary benefit from participation is related to the exchange of information about technology and markets. Management participation in the budget-setting process has been widely studied, and it is thought to have both attitudinal and behavioral consequences. The agency perspective assumes that a significant reason for the existence of participation is the transfer of information between subordinate and superior and that there are potential gains for both parties (e.g., better information, resource allocation, incentive plans, performance, compensation).

Since company budgets are developed by consolidating the divisional plans of the entire company, it is essential to communicate extensively with those in charge of formulating budgets, and coordinate their efforts closely (Ueno & Sekaran, 1992). Involvement in the budgetary process improves manager's understanding of how budget distributions are determined (Wentzel, 2002). And a positive communication climate will strengthen organizational identification, because it is rewarding and thus serves a member's self-enhancement. It invites an employee to participate actively in discussions about organizational issues and involves him or her in decision-making. One may thus categorize oneself more easily as a significant member of an in-group. Moreover, experiencing openness in communication with supervisors can improve initial attitudinal response to budgetary allocations, "a predisposition to support or withhold support of the budget and even to sabotage the budget" (Collins, 1978; Ivancevich & Matteson, 1990) and may add to the employee's feelings of self-worth, because under such conditions she/he will experience being taken seriously. This fair treatment is likely to signify to employees that the organization values and cares about their well being (Smidts, Pruyn, & Cess, 2001). From a social-exchange perspective, this support results in employee reciprocation (Eisenberger, Fasolo, & Davis-LaMastro, 1990). Accordingly, I argue that the opportunity to get involved in and communicate on the budget-setting process increases a manager's feeling of involvement in the budget. Such a feeling increases the manager's commitment to their budget goals.

The a priori discussions cited above suggest that high degrees of communication in the budget-setting process will positively impact his/her budget goal commitment. Hence, the following hypothesis is tested:

H3. Budget communication is positively related to budget goal commitment.

Need for influence arises when an individual wants to have impact (McClelland & Watson, 1973) and "the ability to get things done, to mobilize resources, to get and use whatever it is that a person needs for the goals he/she is attempting to meet" (Kanter, 1977). Evidence suggests that the opportunity for middle-level managers to influence the strategic plan may be limited (Woodridge & Floyd, 1992) because strategic planning has historically been seen as the "preserve" of senior management. While previous research suggests that organizations do benefit if middle managers have the opportunity to influence the strategic plan, it appears that the implementation of the strategic plan continues to be the primary responsibility of this group (Wesley, 1990).

Prior models of budgetary participation suggest that more participation in decision-making is always preferable to less (Locke & Schweiger, 1979). Some research has interpreted these models such that in order for participative budgeting to actually exist, an employee must be able to choose his/her own standard. When decision-control does not exist, the resultant condition is often termed pseudo-participation and the associated benefits of participative budgeting are not expected to materialize. Pasewark and Welker (1990), in a direct application of the Vroom–Yetton model, support a general rule recommending only the highest level of participation (decision-control) to ensure success of participative budgeting, and suggest that participation without influence is worse than no participation because it is demotivating to subordinates.

Arora (1992) notes it is only real participation that is effective, not pseudo-participation. He defines real participation as a situation where an employee has power in determining the outcome of a decision. Pseudo-participation, he adds, occurs when workers have less input and are only made to feel they are participating. Arora concludes that pseudo-participation results in employee frustration and a feeling of powerlessness or failure. Additionally, Hyclak (1987) suggests that participative decision-making programs which offer an employee a voice mechanism (process-control) without any decision-control are a form of pseudo-participation.

The contention of the above argument is that when managers perceive that budget decisions are based on their influence, they should be more likely to commit to a goal because they believe the decision outcome is in line with their expectation. H4 tests for this positive association:

H4. Budget influence is positively related to budget goal commitment.

Budget Goal Commitment and Managerial Performance

Based on goal-setting theory, it is argued that goals affect performance by directing attention and action, mobilizing effort and motivating individuals

to develop goal-attainment strategies (Locke et al., 1981; Locke & Latham, 1990). Commitment to budget goals is particularly important since the productivity of the managers determines, to a large extent, whether the organization is able to achieve its objectives (Wentzel, 2002). Locke, Latham, and Erez (1988) contend that it is virtually axiomatic that if there is no commitment to goals, the goal setting does not work. Numerous studies demonstrate that individuals perform better when they accept and commit to attain a particular goal (Locke & Latham, 1990; Locke et al., 1988). For example, Magner et al. (1996) argued that subordinates who are highly committed to their budget goals seek to "interact with people who can provide insight into their work environment, performance goals, task strategies, and other issues that have an important impact on their performance." Empirical tests tend to support this argument (Wentzel, 2002; Kren, 1990). For example, Kren (1990) found that it is commitment to goal, which acts to mobilize effort and increase persistence and thus is the most direct determinant of performance. Hence, the above argument leads to the following hypothesis:

H5. Budget goal commitment is positively associated with managerial performance.

RESEARCH METHODS

Sample Selection

With the assistance of corporate headquarters from two Fortune 500 firms located in the U.S., data were collected[2] from strategic business unit (SBU) managers.[3] In selecting participants, corporate headquarters were instructed to choose SBU managers based on the following criteria: (1) each participant should have budget responsibility in the subunit; (2) each unit would be an investment center and (3) each manager must have held the position for at least 2 years with the business unit. Questionnaires were distributed to a sample of 237 managers whom corporate headquarters had identified as having budget responsibilities.

A letter of endorsement from corporate headquarters was attached to each questionnaire. In addition, a cover letter to the questionnaire explained the importance of the research and guaranteed the respondents that none of their responses would be disclosed to anybody and that only summary data

from the total responses would be published. To minimize response bias and to enable the respondents to mail these back without the risk even of perusal by secretarial staff, a preaddressed stamped envelop was closed with each questionnaire. I also mailed a second questionnaire, a cover letter ensuring confidentiality and a preaddressed stamped envelop to the supervisors.

Of the 237 questionnaires distributed to managers, respondents returned 181, a response rate of 76.37%. However, eight questionnaires were incomplete. Hence, 173 responses were used in the data analysis.[4]

Measures

Measurement instrument for the variables in the questionnaire were developed from existing studies. The subsequent subsections discuss the instruments and the appendix contains an abbreviated copy of the research questionnaire used to measure the self-reported[5] variables in this study.

Environmental Uncertainty

Environmental uncertainty was measured in terms of the perceived degree of complexity of the various components comprising the firm's environment. The measurement approach is that of Brownell (1985, 1987) and Sathe (1982). Measures of environmental complexity described in terms of the number of elements of an organization's (or subunit's) environment, which are viewed as critical to decision-making. The elements include: (1) the availability of suitable qualified personnel, (2) interdependence with other units within the organization, (3) impact of firm goals and objectives, (4) the demands of product consumers, (5) constraints from suppliers, (6) actions of competitors, (7) impact of government regulation and (8) keeping pace with technological advances. Respondents were asked to rate each element as critical or noncritical for their budget decision-making on a seven-point Likert scale (1 = not critical, 7 = very critical).

Budget Participation

With the exception of Hassel and Cunningham (1993, 1996), previous studies involving budgetary participation (Dunk, 1989; Mia, 1989; Kren, 1992; Magner et al., 1996; Nouri & Parker, 1996) use the six items developed by Milani (1975). These previous studies, though, used a composite measure of the six items based on the premise that they represented a single factor. In this study, similar to Hassel and Cunningham (1996), budget participation is measured using two constructs: budget communication and budget

influence. The following three items measure budget communication on a seven-point Likert scale: (1) frequency of supervisor-initiated budget-related discussions, (2) frequency of manager-initiated budget-related discussion and (3) kind of reasoning provided to managers when the budget is revised. Similarly, budget influence was measured using three items: (1) amount of manager's influence on final budget of the unit, (2) the portion of the budget the manager was involved in setting and (3) The importance of the manager's contribution to the budget. These responses were measured using a seven-point Likert scale (1 = strongly disagree; 7 = strongly agree). The factor analysis of the responses shows that they load on two factors with eigenvalues greater than 1.0 (see Table 1).

Budget Goal Commitment

Budget goal commitment was measured by a seven-point Likert-scale instrument using the following three items based on Latham and Steele (1983) and Erez and Arad (1986): (1) how committed are you to attaining your responsibility area's budget? (1 = not at all committed, 7 = very committed), (2) how important is it to you to at least attain your responsibility area's budget? (1 = very unimportant, 7 = very important) and (3) To what extent

Table 1. Rotated Factor Loadings of Budgetary Participation Items.

Factor Title and Items Loadings	Factor 1 Budget Communication	Factor 2 Budget Influence
1. Frequency of supervisor-initiated budget-related discussions	0.906	
2. Frequency of manager-initiated budget-related discussion	0.805	
3. Kind of reasoning provided to managers when the budget is revised	0.837	
4. Amount of manager's influence on final budget of the unit		0.881
5. The portion of the budget the manager was involved in setting		0.868
6. The importance of the manager's contribution to the budget		0.906

are you striving to attain your responsibility area's budget? (1 = to no extent, 7 = to a great extent). This measure reflects the view of commitment as being an attitude about a goal and the maintenance of that determination. This measure derives commitment levels by asking directly about the goal.

Managerial Performance

The following nine-dimensional self-rating measures were adapted from Mahoney et al. (1963) to evaluate managerial performance on a seven-point Likert scale: (1) planning, (2) investigating, (3) coordinating, (4) evaluating, (5) supervising, (6) staffing, (7) negotiating, (8) representing and (9) rate your overall performance. This measure has been used by prior researchers (Brownell, 1982; Brownell & Hirst, 1986; Dunk, 1989, 1993). Respondents were asked to rate their performance using a seven-point Likert scale (1 = below average; 7 = above average). Brownell (1982) reported that an independent assessment of reliability and validity of the Mahoney instrument have provided supportive evidence of the measure's sound development (Penfield, 1974; Heneman, 1974).

RESULTS

Descriptive Statistics

Information on the characteristics of respondents is provided in Table 2. The mean response to the question regarding the number of years with the business unit in their current position is 6.13 years (S.D. = 1.56).[6] Regarding the length of time in a management position, respondents indicate a mean of 15.42 years (S.D. = 3.17). Given their tenure with the business subunit and their management experience, the respondents are well qualified to provide the information requested. The size of the business subunits are measured as the average number of employees, 347 (S.D. = 53.17), and average business unit sales, $57.579 million (S.D. = 2.347).

Table 2. Respondents' Characteristics.

	Minimum	Maximum	Mean	Standard Deviation
Firm size (number of employees)	229	756	347	53.17
Length at present position (in years)	4.00	17.00	6.13	1.56
Total sales (in million)	23.11	489	57.56	89.47
Length in management (in years)	4.00	26	15.42	3.17

Analysis of the Structural Model

Before evaluating the relationship between environmental uncertainty, budget commitment, budget influence, budget goal commitment and managerial performance, I first assessed the reliability of the measures. Table 3 shows that the reliability measures are 0.750 for environment uncertainty, 0.823 for budget communication, 0.949 for budget influence, 0.872 for budget goal commitment and 0.803 for managerial performance. Hence, all measures demonstrated acceptable reliabilities, with coefficients above 0.70. Taken together, the above results support the use of the constructs and indicator variables for testing the study's hypotheses.

Next, the relationships among the variables of interest are examined using structural equation modeling. Strength of structural equation analysis is that multiple indicators are used to represent each unobserved latent construct and that it provides an efficient technique for estimating interrelated dependence relationships, such as those proposed in this study. The contribution of each scale item is incorporated into the estimation of the independent and dependent relationships of the model. This procedure is similar to performing a factor analysis of the scale items and using the factor scores in a regression analysis.

The overall fit statistics in Table 4 reveal that the proposed model fits reasonably well with the data from the business unit managers. First, the Chi-square (χ^2) test statistic associated with the null hypothesis that the proposed model can effectively reproduce the observed covariance is 500.89 with 251 degrees of freedom resulting in a ratio of 1.99. Good-fitting models evidence ratios of 2.0 or less (Wheaton, Muthen, Alwin, & Summers, 1977). Second, Table 4 shows that the various measures of relative and absolute fit indices (ranging from 0 to 1, with 0 implying poor fit and 1 indicating perfect fit) including the goodness-of-fit indices (GFI), the comparative fit indices (CFI) and the normed fit indices (NFI) exceed 0.90 without any exceptions. Noting that different fit indices have different strengths and weaknesses, this

Table 3. Reliability Measures.

Variables	Reliability
Environmental uncertainty	0.750
Budget communication	0.823
Budget influence	0.949
Budget goal commitment	0.872
Managerial performance	0.803

Table 4. Overall Fit Summary of the Theoretical Model.

	Results	Acceptable Fit Standard
Statistical tests		
Chi-square	500.89	N/A
df	251	N/A
Chi-square/df	1.99	<2.0
Fit indices		
GFI	0.903	>0.90
CFI	0.927	>0.90
NFI	0.916	>0.90
NNFI	0.907	>0.90
Residual analysis		
RMSEA	0.046	<0.05

GFI is the goodness-of-fit index; CFI the comparative-fit-index; NFI the normed-fit-index; NNFI the nonnormed-fit index; RMSEA the root mean square error for approximation.

consistent evidence of exceeding the target value of 0.90 for good-fitting models is encouraging. Third, Table 4 also shows that the difference between reproduced and observed covariances is rather small as evidenced by the root mean square error of approximation (RMSEA) of 0.046. Thus, the proposed model in Fig. 1 is an acceptable and reasonable portrayal of the data and serves as a sound basis for interpreting the specific hypotheses and influence pathways.

Hypotheses Testing

To test the hypotheses, I rely on the standardized parameter estimates for the model (see Table 5 and Fig. 2). Consistent with the theoretical expectations, the standardized parameter estimates between environmental uncertainty and budget communication, and environmental uncertainty and budget influence are positive and significant (path coefficients = 0.385, p = 0.003, 0.331 and 0.002, respectively). Thus H1 and H2, which state that environmental uncertainty has a significant positive impact on budget communication (H1) and budget influence (H2) are supported. Furthermore, the relations between both budget communication and budget influence and budget goal commitment are positive and statistically significant (path coefficients = 0.181, p = 0.029, 0.258 and 0.001, respectively). Thus, H3 and H4 are supported. Further, the impact of budget goal commitment and

Table 5. Estimated Measurement Coefficients.

Standardized Structural Paths	Coefficient	*p*-Value
Environmental uncertainty–budget communication	0.385	0.003
Environmental uncertainty–budget influence	0.331	0.002
Budget communication–budget commitment	0.181	0.029
Budget influence–budget commitment	0.258	0.001
Budget commitment–managerial performance	0.241	0.004

Notes: R^2 for budget communication is 14.80%; R^2 for budget influence is 11%; R^2 for budget goal commitment is 11.10%; R^2 for managerial performance is 6.7%.

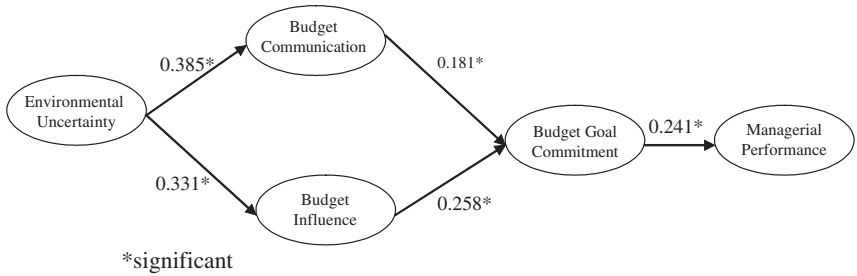


Fig. 2. Model Path Significance Results.

managerial performance is significant and positive (path coefficient = 0.241, $p = 0.004$). Therefore, H5 is supported. Also, the squared multiple correlations (R^2) (Table 5) of the endogenous constructs indicate low-explained variances of 14.80% in budget communication, 11% in budget influence, 11.10% in budget goal commitment and 6.7% in managerial performance.

Further Analyses

Between-model comparisons were undertaken using the χ^2 difference test recommended by Bollen (1989) and others (Hayduk, 1987; Joreskog & Sorbom, 1993; Medsker, Williams, & Holahan, 1994) along with differences in the fit indices (Gerbing & Anderson, 1992; Medsker et al., 1994; Tanaka, 1993). The following models were tested and compared to the theoretical model (model 1) presented in Fig. 1: model 2 tests the relationships between environmental uncertainty and budget goal commitment, model 3 tests the relationship between environmental uncertainty and managerial performance, model 4 tests the relationship between budget communication and

Table 6. Results of Model Comparisons.

Models	χ^2	df	$\Delta\chi^2$	Δdf	GFI	CFI	NFI
1. Theoretical model	500.89	251	—	—	0.829	0.909	0.837
2. Model 2	500.00	250	0.89	1	0.829	0.909	0.837
3. Model 3	500.81	250	0.81	1	0.829	0.909	0.837
4. Model 4	500.70	250	0.19	1	0.829	0.909	0.837
5. Model 5	499.99	250	0.90	1	0.829	0.909	0.837

Notes: Model 2 adds path from environmental uncertainty to budget goal commitment. Model 3 adds path from environmental uncertainty to managerial performance. Model 4 adds path from budget communication to managerial performance. Model 5 adds path from budget influence to managerial performance.

managerial performance and model 5 tests the relationship between budget influence and managerial performance.

Table 6 shows that when each of the models (2–5) are compared to the theoretical model (1), none of the models yields a χ^2 change that is statistically significant. Therefore, on the basis of fit indices, model 1 provides the best fit.

SUMMARY AND DISCUSSION

Using structural equation modeling, the major aim of this study is to investigate the influence of environmental uncertainty on budget communication and budget influence, the impact of both budget communication and budget influence on budget goal commitment and the effect of budget goal commitment on managerial performance. Overall, the results of this study indicate support for the theoretical framework. Environmental uncertainty factors have significant relationships with both budget communication and budget influence which, in turn, have significant positive impact on budget goal commitment. Budget goal commitment has a significant positive influence on managerial performance. Further analyses, using χ^2 differences, support the hypotheses. These findings suggest that environmental uncertainty as an antecedent and budget goal commitment as consequence of both budget participation variables (i.e., budget communication and budget influence) provide information in explaining the relationship between budget participation and managerial performance. This is an important finding as prior studies relating budget participation and managerial performance have been mixed.

While the results for the above hypotheses are believed to be valid and reliable, some limitations of this study should also be noted. First, even though the survey method is a well-documented research methodology, the results of this study may, nevertheless, be affected by the usual limitations associated with generalizability since the sample was obtained from only two Fortune 500 firms. Hence, future research could go well beyond the specific suggestions made here. Second, the low explained variances indicate potential for omitted variables. Therefore, other variables may have more explanatory power than those used in this study. Third, field evidence of these issues and well-designed experimental and archival tests are needed to distinguish among different explanations for observed behavior to support the predictions being tested.

Despite these limitations, the results of this study have several important implications for managers and researchers. The results strongly suggest that the structural equation modeling offers a useful way for managers to approach their budget-setting strategies. In particular, when budget participation is viewed as budget communication and budget influence, both environmental uncertainty and budget goal commitment must be incorporated into development of managerial performance. The variables and the results of this study should enhance practitioners' confidence in their budget design.

This study contributes significantly to the budgeting literature by helping to reconcile the results reported by previous research in this area and thereby improving our understanding of the relationships among factors leading to managerial performance within the context of environmental uncertainty and budget goal commitment. Most interesting in the findings is the strong impact environmental uncertainty has on both budget communication and budget influence which, in turn, significantly influence budget goal commitment, and the positive impact of budget goal commitment on managerial performance. This is an important finding as prior studies relating budget participation and performance have been mixed.

NOTES

1. With one significant exception (Hassel & Cunningham, 1993, 1996), no empirical studies have explicitly addressed budget communication and budget influence separately. This study continues the example of Hassel and Cunningham (1993, 1996) and examines budget communication and influence separately. Because the two dimensions are conceptually independent, it is appropriate to develop hypotheses separately for each dimension.
2. A survey questionnaire was used as a cost-effective method to collect data.

3. The need to obtain access and the constraints of funding prevented the use of random sample from the entire Fortune 500. However, given their size range (in 2002 sales from about 31 to 237 million) and the diversity of industries in which they operate (industrial machinery, electronic components, electronic equipment, instruments and related products, chemicals and primary metal industry), there is no prima facie reason to expect any systematic bias in the findings from business units within these firms.
4. I used discriminant analysis to compare respondents to the first mailing, the early respondents to those responding thereafter, the late respondents (Fowler, 1993). Results revealed that the two groups did not differ significantly in either the level of the variables or in the relationship between the variables at the 0.05 level. This suggests that nonresponse bias may not be a problem.
5. Although self-rating measures have sometimes been criticized for a potential leniency bias, this is less of a concern where such bias is generic and where the ratings are needed for a relative rather than absolute analysis, as is in the present study (Pelham & Wilson, 1996). Also, it has been shown that subjective measures are a reliable method for evaluating performance (Pearce & Robinson, 1987), and research has reported a strong association between objective measures and subjective responses (Robinson & Pearce, 1988; Venkatraman & Ramanujam, 1986).
6. Standard deviation.

REFERENCES

Arora, B. (1992). Participative management or managed participation – the case of India. *Economic and Industrial Democracy*, *13*, 263–272.

Baiman, S., & Evans III, J. H. (1983). Pre-decision information and participative management control systems. *Journal of Accounting Research*, *21*(Autumn), 371–395.

Becker, S., & Gordon, G. (1966). An entrepreneurial theory of formal organizations Part 1. *Administrative Science Quarterly*, 315–344.

Birnberg, J. G., Shields, M. D., & Young, M. Y. (1990). The case for multiple methods in empirical management accounting research (with an illustration from budget setting). *Journal of Management Accounting Research*, 33–66.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Brownell, P. (1981). Participation in budgeting, locus of control and organizational effectiveness. *The Accounting Review*, *56*(4), 844–860.

Brownell, P. (1982). The role of accounting data in performance evaluation, budgetary participation, and organizational effectiveness. *Journal of Accounting Research*, *20*, 12–27.

Brownell, P. (1985). Budgetary systems and the control of functionally differentiated organizational activities. *Journal of Accounting Research*, *23*(Autumn), 502–512.

Brownell, P. (1987). The role of accounting information, environment and management control in multinational organizations. *Accounting and Finance*, *27*, 1–16.

Brownell, P., & Dunk, A. S. (1991). Task uncertainty and its interaction with budgetary participation and budget emphasis: Some methodological issues and empirical investigation. *Accounting, Organizations & Society*, *16*(8), 693–703.

Brownell, P., & Hirst, M. (1986). Reliance on accounting information budgetary participation and task uncertainty: Test of three-way interaction. *Journal of Accounting Research*, *24*, 241–249.

Brownell, P., & McInnes, M. (1986). Budgetary participation, motivation, and managerial performance. *The Accounting Review*, *61*(4), 587–600.

Christensen, J. (1982). The determination of performance standards and participation. *Journal of Accounting Research*, *20*, 589–603.

Collins, F. (1978). The interaction of budget characteristics and personality variables with budgetary response attitudes. *The Accounting Review*, *53*(2), 324–335.

Dunk, A. S. (1989). Budget emphasis, budgetary participation and managerial performance: A note. *Accounting, Organizations and Society*, *14*(4), 321–324.

Dunk, A. S. (1993). The effects of job-related tension on managerial performance in participative budgetary settings. *Accounting, Organizations and Society*, *18*(7/8), 575–585.

Eisenberger, R., Fasolo, P., & Davis-LaMastro, V. (1990). Perceived organizational support and employee diligence, commitment, and innovation. *Journal of Applied Psychology*, *75*(1), 51–59.

Erez, M., & Arad, R. (1986). Participative goal-setting, motivational and cognitive factors. *Journal of Applied Psychology*, *71*(4), 591–597.

Ferris, K. R. A. (1977). Test of the expectancy theory of motivation in an accounting environment: A response. *The Accounting Review*, *52*, 605–615.

Fowler, F. J., Jr. (1993). *Survey research methods*. Newbury Park, CA: Sage.

Gerbing, D. W., & Anderson, J. C. (1992). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. *Sociological Methods and Research*, *21*, 132–160.

Govindarajan, V. (1986). Impact of participation in the budgetary process on managerial attitudes and performance: Universalistic and contingency perspectives. *Decision Sciences*, *17*, 496–516.

Hassel, L., & Cunningham, G. M. (1993). Budget effectiveness in multinational companies: An empirical examination of environmental interaction on cognitive and affective effects of two dimensions of budgetary participation. *Scandinavian Journal of Management*, *9*, 297–316.

Hassel, L., & Cunningham, G. M. (1996). Budget effectiveness in multidimensional corporations: An empirical test of the use of budget controls moderated by two dimensions of budgetary participation under high and low environmental dynamics. *Management International Review*, *36*(3), 245–266.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essentials and advances*. Baltimore, MD: John Hopkins Press.

Heneman, H. G. (1974). Comparisons of self and superior ratings of managerial performance. *Journal of Applied Psychology*, *59*(5), 638–642.

Hofstede, G. H. (1968). *The game of budget control*. London, UK: Assen Tavistock.

Hopwood, A. G. (1976). *Accounting and human behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Hyclak, T. (1987). Worker self-management and economic reform in Poland. *International Journal of Social Economics*, *14*(7–9), 127–135.

Ibrahim, M. M. (1993). The perception of uncertainty in budgetary decision making. *The International Journal of Public Sector Management*, *6*(3), 31–37.

Ivancevich, J. M., & Matteson, M. T. (1990). *Organizational behavior and management* (2nd ed.). Homewood, IL: Irwin.

Joreskog, K. G., & Sorbom, D. (1993). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software.

Kanter, R. M. (1977). *Men and women of the corporation*. New York, NY: Basic Books.

Kenis, L. (1979). Effects of budgetary goal characteristics on managerial attitudes and performance. *The Accounting Review*, *54*(4), 707–721.

Kerby, A. J., Reichelstein, S., Sen, P. K., & Piak, T. Y. (1991). Participation, slack, and budget-based performance evaluation. *Journal of Accounting Research*, *29*(1), 109–128.

Kren, L. (1990). Performance-based control system: An extended expectancy theory model approach. *Journal of Management Accounting Research*, *2*, 100–112.

Kren, L. (1992). Budgetary participation and managerial performance: The impact of information and environmental volatility. *The Accounting Review*, *67*(July), 511–526.

Latham, G. P., & Steele, T. P. (1983). The motivational effects of participation versus goal-setting on performance. *Academy of Management*, *26*, 406–417.

Lawrence, P. R., & Lorsch, F. W. (1967). *Organization and environment*. Boston, MA: Harvard University.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ: Prentice-Hall.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin*, *90*, 125–152.

Locke, E. A., Latham, G. P., & Erez, M. (1988). The determinants of goal commitment. *Academy of Management Review*, *13*(1), 23–39.

Magner, N., Welker, R. B., & Campbell, T. L. (1996). Testing a model of cognitive budgetary participation processes in a latent variable structural equations framework. *Accounting and Business Research*, *27*(1), 41–50.

Mahoney, T. A., Jerdee, T. H., & Carroll, S. J. (1963). *Development of managerial performance: A research approach*. Cincinnati, OH: Southwestern Publishing Company.

McClelland, D. C., & Watson, R. I., Jr. (1973). Power motivation and risk-taking behavior. In: D. C. McClelland & R. S. Steele (Eds), *Human motivation: A book of readings*. Morristown, NJ: General Learning Press.

Medsker, G. J., Williams, L. J., & Holahan, P. J. (1994). A review of current practice for evaluating causal models in organizational behavior and human resources management research. *Journal of Management*, *20*, 439–464.

Merchant, K. A. (1981). The design of the corporate budgeting system: Influences on managerial behavior and performance. *The Accounting Review*, *56*(October), 813–829.

Mia, L. (1988). Managerial attitude, motivation and the effectiveness of budget participation. *Accounting, Organizations and Society*, *13*(5), 465–475.

Mia, L. (1989). The impact of participation in budgeting and job difficulty on managerial performance and work motivation: A research note. *Accounting, Organizations and Society*, *14*, 347–357.

Milani, K. (1975). The relationship of participation in budget setting to industrial supervisor performance and attitudes: A field study. *The Accounting Review*, *50*(April), 274–284.

Miller, K. I., & Monge, P. R. (1986). Participation, satisfaction, and productivity. *Academy of Management Journal*, *25*(4), 727–753.

Murray, D. (1990). The performance effects of participative budgeting: An integration of intervening and moderating variables. *Behavioral Research in Accounting*, *2*, 104–123.

Nouri, H., & Parker, R. J. (1996). The effect of organizational commitment on the relation between budgetary participation and budgetary slack. *Behavioral Research in Accounting*, *8*, 74–90.

Pasewark, W., & Welker, R. (1990). A Vroom–Yetton evaluation of subordinate participation in budgetary decision making. *Journal of Management Accounting Research*, *2*, 113–126.

Pearce, J. A., II, & Robinson, R. B., Jr. (1987). A measure of CEO social power in strategic decision-making. *Strategic Management Journal*, *8*(3), 297–304.

Pelham, A. M., & Wilson, D. T. (1996). A longitudinal study of the impact of market structure, firm structure, strategy, and market orientation culture on dimensions of small-firm performance. *Academy of Marketing Science Journal*, *24*(1), 27–43.

Penfield, R. V. (1974). Time allocation patterns and effectiveness of managers. *Personnel Psychology*, *27*(2), 245–255.

Penno, M. (1990). Accounting systems, participation in budgeting, and performance evaluation. *The Accounting Review*, *65*(2), 303–314.

Robinson, R. B., Jr., & Pearce, J. A. (1988). Planned patterns of strategic behavior and their relationships. *Strategic Management Journal*, *9*(1), 43–60.

Sathe, V. (1982). *Controller involvement in management*. Englewood Cliffs, NJ: Prentice Hall.

Schweiger, D. M. (1979). Participation in decision-making: One more look. *Research in Organizational Behavior*, *1*, 265–339.

Shields, J. F., & Shields, M. D. (1998). Antecedents of participative budgeting. *Accounting, Organizations and Society*, *23*, 49–76.

Shields, M., & Young, S. M. (1993). Antecedents and consequences of participative budgeting: Evidence on the effects of information asymmetry. *Journal of Management Accounting Research*, *5*, 265–280.

Simons, R. (1987). Accounting control systems and business strategy: An empirical analysis. *Accounting, Organizations and Society*, *12*(4), 357–374.

Smidts, A., Pruyn, A. T. H., & Cess, B. M. V. R. (2001). The impact of employee communication and perceived external prestige on organizational identification. *Academy of Management Journal*, *44*(5), 1051–1062.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In: K. A. Bollen & J. S. Long (Eds), *Testing structural equation models*. Newbury Park, CA: Sage.

Tichy, M. (1982). Managing change strategically: The technical, political, and cultural keys. *Organizational Dynamics*, *11*(Autumn), 59–80.

Ueno, S., & Sekaran, U. (1992). The influence of culture on budget control practices in the U.S.A. and Japan: An empirical study. *Journal of International Business Studies*, *23*(4), 659–674.

Venkatraman, N., & Ramanujam, V. (1986). Measurement of business performance in strategy research: A comparison of approaches. *The Academy of Management Review*, *11*(4), 801–814.

Vroom, V. H., & Jago, A. G. (1988). *The new leadership: Managing participation in organizations*. Englewood Cliffs, NJ: Prentice-Hall.

Wentzel, K. (2002). The influence of fairness perceptions and goal commitment on managers' performance in a budget setting. *Behavioral Research in Accounting*, *14*, 247–271.

Wheaton, B., Muthen, B., Alwin, D., & Summers, G. (1977). Assessing reliability and stability in panel models. In: D. Heise (Ed.), *Sociological methodology* (pp. 84–136). San Francisco, CA: Jossey-Bass.

Woodridge, B., & Floyd, S. W. (1992). Middle management involvement in strategy and its association with strategic type: A research note. *Strategic Management Journal*, *13*, 153–167.

APPENDIX

I. Environmental Uncertainty

Please rate the following in terms of their criticality (1 = not critical, 7 = very critical)

1. The availability of suitable qualified personnel
2. Interdependence with other units within the organization
3. Impact of firm goals and objectives
4. The demands of product consumers
5. Constraints from suppliers
6. Actions of competitors
7. Impact of government regulation
8. Keeping pace with technological advances.

II. Budget Participation

Budget communication (1 = strongly disagree, 7 = strongly agree)

1. Frequency of supervisor-initiated budget-related discussions
2. Frequency of manager-initiated budget-related discussion
3. Kind of reasoning provided to managers when the budget is revised.

Budget influence (1 = strongly disagree, 7 = strongly agree)

1. Amount of manager's influence on final budget of the unit
2. The portion of the budget the manager was involved in setting
3. The importance of the manager's contribution to the budget.

III. Budget Goal Commitment

1. Commitment to a goal means acceptance of it as your personal goal and your determination to attain it. How committed are you to attaining your responsibility area's budget? (1 = not at all committed, 7 = very committed)
2. How important is it to you to at least attain your responsibility area's budget? (1 = very unimportant, 7 = very important)
3. To what extent are you striving to attain your responsibility area's budget? (1 = to no extent, 7 = to a great extent).

IV. Managerial Performance

Please rate your performance on the following tasks: (1 = below average; 7 = above average)

1. Planning
2. Investigating
3. Coordinating
4. Evaluating
5. Supervising
6. Staffing
7. Negotiating
8. Representing
9. Rate your overall performance.

This page intentionally left blank

THE IMPACT OF EMPLOYEE RANK ON THE RELATIONSHIP BETWEEN ATTITUDES, MOTIVATION, AND PERFORMANCE

Stan Davis and James M. Kohlmeyer, III

ABSTRACT

In this paper, we investigate the effect that employee rank has on attitudes and performance where supervisors establish budgeted standards of performance. This paper advances the extant management accounting literature by considering a variable (employee rank) not considered in prior related studies. Our findings indicate the impact of attitudes on performance is moderated by the rank of the employee within the organization. We find lower ranked employees within the organization performed better when they felt the process for establishing their performance standards was fair. For employees in higher ranking positions, the motivation associated with feedback on their performance was a factor in determining performance, while the degree to which they felt the process for establishing standards was fair was not.

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 233–252

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14011-8

INTRODUCTION

Previous research (Chenhall & Brownell, 1988; Fisher, Frederickson, & Peffer, 2002; Lindquist, 1995; Nouri & Parker, 1996) has explored the effects of employee attitudes on performance in settings where subordinates have varying degrees of input into their budgeted standards of performance. These studies build support for a theoretical model that suggests when employees are satisfied with the process associated with the establishment of their budgeted standards of performance, they are more satisfied with the budget and ultimately perform better. The purpose of this paper is to extend this line of research by considering a potentially moderating variable that has heretofore been ignored in the accounting literature: employee rank. Our goal is to determine if previously observed relationships between attitudes, motivation, and performance hold across employees at differing ranks within the organization. This study is motivated by the continuing need for accounting researchers and managers to understand the factors that affect employee performance, especially in performance standard setting environments.

We rely on survey and archival data from the field site to conduct this study. The setting is a mid-sized southeastern US bank that recently implemented a balanced scorecard (BSC) program to direct and evaluate its employees' performance. Survey data was collected during the first year of the BSC program. Archival data were also collected from the field site for the corresponding time period.

We find that an employee's rank within the organization moderates the relationships between attitudes, motivation, and performance. Specifically, we find that lower ranking employees within the organization perform better when they were in agreement with their performance standards established by their supervisor. For employees in higher ranking positions, the motivation associated with feedback on their performance was a factor in determining performance, while whether they were in agreement with their standards was not a factor in determining their performance level.

This study highlights the importance of fairly perceived standard-setting procedures and broadens our understanding of the relationship between employee attitudes, motivation, and performance by observing differences in motivating factors between employees at different ranks or levels within the organization. The study's findings also provide accounting researchers with another variable of consideration as they continue to explore factors that affect individual performance within organizations.

The remainder of this paper is organized as follows. The relevant literature and theory is discussed in the next section along with the presentation of hypotheses. The third section explains the research method and the fourth section discusses the results. The concluding section discusses the implications of findings, the study's limitations, and directions for future research.

LITERATURE REVIEW AND HYPOTHESIS FORMULATION

The Relationship between Attitudes and Performance in Budget Setting Contexts

Two recent studies have investigated the relationship between attitudes and performance in a budget setting context. In an experimental study involving undergraduate students, Lindquist (1995) found that subordinates with a voice in the establishment of their performance standards were more satisfied with the budget and the task to be performed, even in situations where the established standards of performance were thought to be unfair by the subordinate.[1] This study did not, however, find support for the hypotheses that predicted greater levels of performance when subordinates had a voice in the budget setting process.

Fisher et al. (2002) recently established support for a theoretical model based on procedural justice theory and goal setting theory that links negotiated agreement with performance standards to subordinate performance. Negotiated agreement is linked to subordinate performance through degree of process control, perception of budgeting process, perception of budget, and budget commitment. In an experiment using undergraduate accounting students, they find subordinates who reach a negotiated agreement on their budgeted standards of performance perceive the process for establishing budgets as fair, are satisfied with their budget, are committed to achieving the budget, and perform at higher levels than those that do not reach a negotiated agreement.

The goal setting literature shows that commitment to a budget is higher when subordinates perceive the budget as coming from a legitimate source (Locke & Latham, 1990). When subordinates perceive the budget-setting process to be unfair, they are less likely to view the budget as legitimate. A central tenant of goal setting theory is that for a given goal, individuals who have higher commitment to the goal exert higher effort, whereas in-

dividuals who have lower commitment exert lower effort (Hollenbeck & Klein, 1987; Latham, 2004; Lee, Locke, & Phan, 1997; Locke & Latham, 1990). Thus, a positive relationship between commitment to budgeted levels of performance and subordinate performance also should exist.

With the budget-setting process analogous to the setting of BSC targeted measures of performance, we apply the goal setting literature and previous research findings to this setting. The previous discussion suggests that the attitudes and performance of subordinates who fail to agree with their established performance budgets should differ from those who do agree. Specifically, subordinates who do not agree with their established targeted levels of performance will be less satisfied with their budgets, less committed to attaining their budgets, and perform at lower levels.[2]

The Effect of Feedback Motivation on Performance

Procedural fairness perceptions alone do not account for commitment to and achievement of performance goals. Goal-setting research is replete with theory and evidence that shows motivation also affects employee performance (Locke & Latham, 1990). Further, feedback has been shown as a means to improve performance from a cognitive and motivational standpoint.[3] In a review of research on objective feedback, Kopelman (1986) observes that feedback results in improved task performance in part because it increases motivation and leads to the anticipation of gaining (or losing) external rewards. Where feedback is perceived as being meaningful and important, the expectation is that greater performance will ensue.

The Role of Employee Rank on Justice Perceptions

Recent research on fairness perceptions suggests that employees at different hierarchical levels within the organization may be affected by justice perceptions differently (Schminke, Cropanzano, & Rupp, 2002; Hunton, Hall, & Price, 1998). Homans' (1974) work has important implications for the relationship between social exchange at different organizational levels and justice perceptions. Relative to lower level employees, those who are higher in the organizational hierarchy tend to experience higher levels of distributive and procedural elements like more pay, more influence over policies, and being treated with greater respect (Aquino, Grover, Bradfield, & Allen, 1999; Brass & Burkhardt, 1993; Finkelstein & Hambrick, 1990). Schminke

et al. (2002) contend that the positive impact of higher justice perceptions is less for employees with higher ranking positions in the organization as compared to employees in lower ranking positions. Using work from Homans (1974) as a basis for their theory building, Schminke et al. (2002) argue that incremental units of a desired resource are valued more greatly by individuals with initially lower levels of that resource than by individuals with initially greater levels of that resource. Placing this concept in the domain of employees in an organization dealing with justice issues, they state that "providing additional voice to a high ranking official who already has a great deal of authority to establish procedures may make only a small impression. But consulting the opinion of a lower level worker, who might otherwise feel alienated, disenfranchised, and voiceless, could have profound ramifications" (Schminke et al., 2002, p. 886).

Schminke et al. (2002) find support for the moderating effects of employee rank on the impact of various organizational structural characteristics. Specifically, they find that the justice perceptions of employees in lower ranking positions generally are affected by organizational structural characteristics of centralization (participation and authority hierarchy), formalization, complexity and size more so than employees in higher ranking positions. For higher ranking employees, they posit that justice perceptions arise less from organizational structural characteristics and more from the broader benefits associated with being higher ranked, primarily the experiencing of higher quality social exchange relationships. Schminke et al. (2002) surmise that for situations where justice perceptions may be affected, the sense of unfairness tends to fall more heavily on those with less power, those in lower ranking positions. Higher ranking employees tend to have more positive experiences within the organization, thereby lessening the impact of unfair outcomes when they experience them. They conclude that employee rank within the organization is a valid and important factor to consider when studying justice perceptions.

Consistent with expectations set forth by Schminke et al. (2002), Hunton et al. (1998) found that an increase from a no-voice position to a low level of voice resulted in an increase in perceived fairness for student subjects; however, similar marginal increases in voice did not result in a change in perceived fairness for executives. Hunton et al. (1998) found that much larger increases in voice were necessary before executives perceived greater fairness in the same setting, reinforcing Schminke et al.'s (2002) argument that those without a valuable resource are impacted more than those already with the resource when receiving an incremental unit of the valued resource. Placing this finding in the context of the current study, we contend that lower

ranking employees that experience voice in the budgeting process will be more positively impacted by the perceived fairness associated with having a voice in the process than higher ranking employees who already experience multiple fair outcomes from the organization.

Hypothesis Development

Results from the prior research discussed above indicate that fairness perceptions (as they relate to the budget-setting process) affect satisfaction, commitment, and performance (Lindquist, 1995; Fisher et al., 2002). Further, prior research has indicated that feedback motivation is an important factor that can lead to improved performance (see Kopelman, 1986, for example). We therefore begin our analysis by testing for these findings in a different research setting than an experimental lab, using employees in an organization that had recently implemented a BSC as the subjects for the study. The budget-setting context for this setting is the establishment of performance targets on their individual BSCs. Given this setting, we offer the following hypotheses:

H1a–c. Employees that are in agreement with their budgeted standards of performance will (a) report higher levels of satisfaction with the budget, (b) report higher levels of commitment to the budget, and (c) perform at higher levels than employees that are not in agreement with their budgeted standards of performance.

H2. Employees that report being motivated by the feedback they receive on their performance will perform at higher levels than employees that report not being motivated by the feedback they receive on their performance.

As discussed above, recent research has suggested that the employee's ranking within the organization may have an affect on the development and strength of justice perceptions for employees at different ranks within the organization. Schminke et al. (2002) found that in situations where decision-making was centralized, employees in lower ranking positions experienced lower levels of distributive justice as compared to employees in higher ranking positions. We therefore explore the previous two hypotheses within two subgroups of employees: lower ranking (LOWRANK) employees and higher ranking (HIGHRANK) employees. Based on the findings of Schminke et al. (2002) and Hunton et al. (1998), LOWRANK employees

tend to be more keenly aware and affected by justice issues due, in part, to their lack of power; therefore, we expect that LOWRANK employees in agreement with budgeted standards of performance, a distributive justice outcome measure, will respond favorably to a just outcome and be more satisfied, more committed, and perform at a higher level than LOWRANK employees not in agreement (due to the positive effects of having experienced a just outcome of the budget-setting process). Further, we do not necessarily expect the same for HIGHRANK employees. Since HIGHRANK employees tend to already experience just outcomes on a more regular basis, the presence of a just outcome (perceived fair budgeted levels of performance) in this setting should not necessarily provide the same positive benefits it does for LOWRANK employees. Implicit in this argument is the contention that for HIGHRANK employees, having a voice in the establishment of their budgeted levels of performance is just one of several fair exchanges these employees have with the organization. Therefore, we expect employee rank to moderate the hypothesized relationships set forth in H1a–c, and accordingly, we set forth the following hypothesis:

> **H3a–c.** The relationship between "agreement with budgeted standards of performance" and level of satisfaction with the budget, (a) level of commitment to the budget, and (c) performance will be moderated by employee rank.

Finally, after examining the moderating effects of employee rank on the relationship between budget agreement, budget satisfaction, budget commitment, and budget performance, we explore the moderating effect of employee rank on the other motivational factor of interest in this study: feedback motivation. Applications of social exchange theory and expectancy theory lend support for expecting differing motivational effects of feedback on high- versus low-ranking employees.

Close social exchange relationships involve open-ended trust and identification with the organization (Rousseau, 1995), with these types of relationships more likely to exist at higher ranks in the organization, where individuals receive greater rewards and tend to have a stronger role in shaping policy than at lower ranks (Schminke et al., 2002). Further, research suggests that higher ranked employees tend to have greater organizational commitment (Mathieu & Zajac, 1990), are more involved in their work (Brown, 1996) and are more apt to have a long-term horizon (Chen, 1995; Martin & Harder, 1994) than employees at lower ranks in the organization. Given these findings, we contend that higher ranked employees will value meaningful feedback more than lower ranked employees due to their greater

association with the organization (i.e., greater commitment, involvement, and long-term perspective). Because feedback represents a reflection of what the organization views as important (as it relates to the employee), it serves as an influence to perform at greater levels for employees more closely associated with the organization. Whereas lower ranked employees also may perceive feedback as meaningful, it is less likely that such feedback will provide the same level of motivation to perform at greater levels because these employees tend to be less associated with the organization.

Expectancy theory (Vroom, 1964) also supports the notion that higher ranked employees are motivated more by meaningful feedback than are lower ranked employees. Given that an individual's motivation is affected by the valence (attractiveness) of the outcome, higher ranked employees may see attending to meaningful feedback as a means by which to attain more of the positive outcomes available to them (that are not generally available to employees at lower ranks), including salaries, bonuses, or other forms of compensation or perquisites.

Based on the above arguments, we expect that employee rank will moderate the relationship between feedback motivation and employee performance. Thus, the following hypothesis is presented.

H4. The relationship between feedback motivation and performance will be moderated by employee rank.

METHOD

Field Site Description

The field site for this study is a mid-sized community bank (''the bank'') located in the southeast United States. The bank employs approximately 480 employees at over 45 branches and one headquarters location; all bank locations are located within one state. The branches are grouped into 14 banking centers and these banking centers are divided into three geographic regions. The bank produces internal financial reports at the banking center, region, and bank-wide level on a monthly basis.

The bank began implementation of a BSC program in 2001 with communication and education taking place between top-level bank employees. Throughout 2002, mid- and low-level managers consulted with and developed BSCs for bank employees. Each employee's scorecard included objective measures and targeted levels (or standards) of performance.

According to bank management, however, the actual measures and targeted levels of performance were determined by the supervisors, with input from subordinates. The bank began collecting BSC data and reporting individual performance results to employees in 2003.

Data Collection

We conducted an internet web-based survey of bank employees during the seventh month of the BSC program being in effect. The survey gathered information pertaining to employee attitudes and behaviors as well as demographic information. Respondents were asked to identify themselves on the survey as a means of matching responses to individual BSC performance data collected from the bank. Bank employees received an email from an executive vice-president of the bank, asking them to participate in the survey. The email contained a link to a website that hosted the survey; participating employees had their names placed in a drawing for one of 10 cash prizes of $50. The website containing the survey was hosted by one of the researchers, thereby ensuring respondents' confidentiality.

The survey and subsequent performance data collection generated 155 complete responses out of 368 total responses. We discarded responses for incomplete data (eight responses) or in cases where the employee indicated the belief that their supervisor was unaware of their performance capability (58 responses).[4] While a large majority of responses (77%) were from females, this was not appreciably different from the bank's overall female population of 81%.

Budget-Setting Process at the Field Site

The process to establish budgeted levels of performance can take several different forms, ranging from no participation or negotiation (resulting in either the subordinate or supervisor selecting budgeted standards of employee performance) to full formal negotiations. The field site for this study did not utilize a formal established process for establishing employee performance standards on the BSC. Rather, supervisors consulted with subordinates (to varying degrees across different branches) and then established performance standards for their subordinates.

Attitudinal Measures

We collected single-item attitudinal data as it related to the setting of performance expectations in the BSC program. The bank required a parsimonious survey. While validated multi-item measures would have been preferable, we adopted questions very similar in wording to the single-item measures used by Fisher et al. (2002). Specifically, we asked participants to respond to the following questions:[5] "The performance targets for me on my 'StatSheet' were set at an achievable level", "I am satisfied with my final 'StatSheet'", "I am very committed to attaining the targeted levels of performance on my 'Statsheet'", and "Is the feedback associated with your performance on the 'Statsheet' meaningful?". Bank officials were very concerned with the wording of this "motivation" question, and as such, a compromise was reached. We associate meaningful feedback to being motivated by feedback in this study and recognize this as a potential limitation to the interpretations of our findings. These questions were not pretested with employees of the bank. For our comparison tests, we segment the responses into HIGH AGREE and LOW AGREE groups based on their response to question, "The performance targets for me on my 'StatSheet' were set at an achievable level", with responses from 1 to 4 being classified as LOW and responses 5–7 being classified as HIGH.[6] We then compare the satisfaction, commitment, and performance of the two groups to determine if a significant difference exists between each group.

Performance Measure

An employee's performance was determined using their actual score on their BSC as of the end of the seventh month of being on the program. Employee scorecards were designed so that their aggregate performance on multiple measures is captured in one number that can range from 0 to 200. Scorecard points are earned as employee's progress toward targeted performance levels on various measures. Performance levels or expectations were established by supervisors with the goal of being able to compare overall performance between any two employees by comparing their overall score on their scorecards. Therefore, a bank teller achieving a score of 125 on her scorecard is perceived by the bank to have performed equally as well as a loan officer receiving the same score.

The task of equating performance standards ratings between employees and job categories was difficult and undoubtedly not achieved perfectly. To

the extent that errors in the setting of performance measurement standards are unsystematic between job titles or levels, we contend that using this measure of performance is valid. However, two important caveats bear discussing. First, if systematic biases exist in the measurement of performance, then the ability to draw conclusions from the findings is somewhat lessened. For example, measures appearing on a lower level employee's scorecard (a teller, e.g.) differ from measures appearing on a higher level employee's scorecard (a branch president, e.g.) such that it may simply be easier to achieve a higher score on your scorecard if you are a lower level employee as compared to a higher level employee, irrespective of your level of satisfaction, commitment, or even effort toward the goal itself. Second, the possibility exists that for employees in disagreement with the standards set for them, the standards were indeed set at impossible levels. Although scorecard designers at the bank took great care to avoid such issues, to the extent that we cannot rule out the possibility of such biases in performance measurement, our results should be interpreted with care.[7]

RESULTS

Hypotheses 1 and 2

Table 1 presents descriptive statistics and correlations for budget agreement, budget satisfaction, budget commitment, and performance. The first set of hypotheses (H1a–c) states that employees in agreement with the level of

Table 1. Descriptive Statistics and Correlations for Attitudinal Variables and Performance Measure.

Measure (n = 155)	Mean	S.D.	1	2	3
1. Agreement to budget	5.09	1.37			
2. Budget satisfaction	4.91	1.48	0.77*		
3. Budget commitment	5.92	1.03	0.517*	0.553*	
4. Performance measure	116.93	27.70	0.10	0.11	0.193*

Performance was measured by an employee's actual score on their balanced scorecard as of the end of the seventh month of being on the program, with possible scores ranging from 0 to 200. Items 1–3 were measured on a seven-point scale anchored by "Strongly Disagree" and "Strongly Agree".

*Correlation is significant at the 0.05 level (2-tailed).

performance standards set for them by a superior will have significantly higher budget satisfaction, commitment, and performance measures than employees that do not agree with their performance standards. The second hypothesis states employees that reported being motivated by the feedback received concerning their performance will outperform employees that are not motivated by the feedback received. Table 2 presents the results of the hypothesis testing.

The findings shown in Table 2 provide partial support for hypotheses 1 and 2. Results in Panel A of Table 2 show employees in the HIGH AGREE group were more satisfied with the BSC program (H1a; $p < 0.01$), and were more committed to attaining their performance standards (H1b; $p < 0.01$) as compared to employees in the LOW AGREE group. However, HIGH AGREE employees had only marginally higher performance than LOW AGREE employees (H1c; $p < 0.10$).

Table 2. Comparison of Employee Attitudes and Performance segmented by agreement with performance standards level and feedback motivation.

Panel A: Descriptive statistics and tests for the effect of 'agreement with standards'

	Agreement with Standards			
	High	Low	*t*-stat	*p*-value*
N	105	50		
Budget Satisfaction (H1a) mean (sd)	5.21 (1.31)	3.56 (1.25)	7.426	<0.01
Budget commitment (H1b) mean (sd)	6.26 (0.64)	5.22 (1.31)	6.635	<0.01
Performance (H1c) mean (sd)	119.37 (26.37)	111.82 (29.95)	1.594	<0.10

Panel B: Comparison of Employee Performance segmented by level of motivation

	Motivation Level			
	High	Low	*t*-stat	*p*-value*
N	103	49		
Performance on BSC	120.84 (25.16)	110.10 (31.53)	2.088	<0.03

*Directional expectation; one-tail *p*-value.

Attitudinal scale responses for the current study were taken on a seven-point scale.

The second hypothesis relies on theory that indicates employees motivated by feedback outperform employees who are less motivated by feedback. We first separate responses into two groups based on the degree to which they are motivated by the feedback received on their BSC and compare performance between the groups; we expect the more highly motivated group to outperform the less motivated group.

Panel B of Table 2 presents the results of a *t*-test comparing the performances of employees reporting to be more highly motivated by feedback to the performance of less motivated employees. Employees reporting to be more highly motivated by the feedback associated with their StatSheet performance outperformed employees reporting to be less motivated by StatSheet feedback (H2; $p < 0.03$).

Hypothesis 3 and 4

Findings presented above indicate both agreement with the budget and feedback motivation can predict greater performance. However, recent research on fairness perceptions suggests that employees at different hierarchical levels within the organization may be affected by justice perceptions differently (Schminke et al., 2002). As such, in an effort to further the understanding of how attitudes, motivation, and performance are related, we separate the survey respondents into two groups based on their job ranking within the organization. Lower ranking employees are those with job titles of teller, customer service representative, administrative assistant, or loan assistant. Higher ranking employees are those with job titles of loan officer, branch management, nonbranch management, and executive management.[8] We repeat H1a–c on each subgroup of employees separately. Hypothesis 3 contends that the relationship between agreement, satisfaction, commitment, and performance relative to the budget is moderated by employee rank. The results for these tests are found in Table 3.

Findings shown in Panel A of Table 3 provide only partial support for H3a–c. For budget satisfaction and budget commitment, both LOWRANK and HIGHRANK employees in the HIGH AGREE grouping were more satisfied with their overall budget (H3a; $p < 0.01$ for both ranks of employees), and were more committed to attaining their performance standards (H3b; $p < 0.01$ for both ranks of employees) as compared with employees in the LOW AGREE grouping. However, for LOWRANK employees, the HIGH AGREE employees significantly outperformed LOW AGREE employees (H3c; $p < 0.05$) while for HIGHRANK employees, no significant

Table 3. Tests performed on LOWRANK and HIGHRANK employee subgroups.

Panel A: Comparison of Employee Attitudes and Performance segmented by agreement with performance standards levels

	Results for LOWRANK employees				Results for HIGHRANK employees			
	Agreement with Standards				Agreement with Standards			
	High	Low	*t*-stat	*p*-value*	High	Low	*t*-stat	*p*-value*
N	56	30			49	19		
Budget satisfaction mean (sd)	5.16 (1.39)	3.43 (1.36)	5.551	<0.01	5.27 (1.24)	3.74 (1.10)	4.708	<0.01
Budget commitment mean (sd)	6.27 (0.62)	5.10 (1.47)	5.172	<0.01	6.24 (0.66)	5.37 (1.07)	4.091	<0.01
Performance mean (sd)	124.77 (29.28)	112.71 (34.14)	1.987	<0.05	113.20 (21.24)	111.22 (23.45)	0.335	<.30

Panel B: Comparison of Employee Performance segmented by level of motivation

	Results for LOWRANK employees				Results for HIGHRANK employees			
	Motivation Level				Motivation Level			
	High	Low	*t*-stat	*p*-value*	High	Low	*t*-stat	*p*-value*
N	50	33			52	16		
Performance on BSC	123.53 (29.92)	118.44 (33.85)	0.718	>.40	118.72 (19.51)	92.90 (16.38)	4.793	<0.01

*Directional expectation; one tail *p*-value

performance difference existed between LOW AGREE and HIGH AGREE employees ($p > 0.30$). Therefore, employee rank does not appear to moderate the relationship between budget agreement and budget satisfaction, or budget commitment but does appear to moderate the relationship between budget agreement and performance, providing support for only H3c.

Panel B of Table 3 presents results of comparison tests for H4, where we expected to see employee rank moderate the relationship between feedback motivation and performance. For LOWRANK employees, employees reporting to be more highly motivated by feedback performed at essentially the same level as employees reporting to be less motivated by feedback ($p > 0.20$). This finding differs from the finding above for H2 using the entire set of respondents.

For HIGHRANK employees, employees reporting to be more highly motivated by feedback had significantly higher performance than employees reporting to be less motivated by feedback ($p < 0.01$). This finding is the same as the finding reported using the whole dataset, but differs from the finding reported using only the LOWRANK employees' responses.

DISCUSSION AND CONCLUSION

In this study, we sought to extend previous research on the relationship between employee attitudes, motivation, and performance in a budget-setting context. Previous research has studied similar issues in laboratory experiments without considering the effects of employee rank. We first replicate similar studies in a field setting (a banking organization) that set budgeted performance standards for employees as they went through the development and implementation of a BSC. We then considered another variable (employee rank) that potentially could lend more insights into our understanding of the relationship between attitudes, motivation, and performance.

Similar to previous studies (Lindquist, 1995; Fisher et al., 2002), we find differences in satisfaction and commitment between employees that agreed with their established performance standards and employees that did not agree. However, we find that the hypothesized relationship between "agreement with standards" and performance is dependent on the ranking of the employee within the organization. We find a difference in performance based on whether an employee agrees with their standards only for employees in lower ranking positions. We find no such difference in performance based on agreement with standards when using HIGHRANK

employees as a subject base. These findings support the contention that fairness perceptions tend to have a stronger impact on LOWRANK employees than they do on HIGHRANK employees, and are consistent with the speculation of Schminke et al. (2002).

Further, we find that effects of feedback motivation are different between low ranking and high-ranking employees. For LOWRANK employees, we found no difference in performance between those that indicated they were motivated by the feedback received pertaining to their performance and those that indicated they were not motivated by the feedback received. For HIGHRANK employees, however, we found a significant difference in performance based on the motivation level associated with the feedback received pertaining to their performance. This finding, coupled with the previous findings, is interesting in that justice perceptions appear to be a better predictor of performance for LOWRANK employees, while feedback motivation appears to be a better predictor of performance for HIGHRANK employees.

These findings have implications both for accounting researchers and practitioners as they continue to understand budget-setting practices, motivation, and performance issues. We confirmed previous experimental findings in a field setting, while identifying a new moderating variable of interest (rank of employee within the organization) for accounting researchers to consider. When not considering the rank of employees within the organization, we found that "agreement with budget standards" proved only a weak indicator of performance. Further, we found that the motivation provided by meaningful feedback proved an indicator of performance for HIGHRANK employees but not for LOWRANK employees. We believe these findings highlight the value of examining the results of experimental studies using field data not only for confirmation of previous findings, but also because extensions offer the opportunity to study other variables of interest, which are more difficult to examine in the laboratory. The final implication for accounting researchers is the observation that care must be used when using students as proxies for employees on tasks where common workplace attitudes or perceptions are observed. Considering the results of previous experimental studies of Lindquist (1995) and Fisher et al. (2002) and the current study, it could be concluded that the students participating in the experimental studies adopted the attitudes of LOWRANK employees. If students regularly adopt the attitudes of LOWRANK employees in similar tasks that may have implications for drawing conclusions from studies on attitudes and performance where students are used as proxies for employees in an organizational setting.

For practitioners, our findings highlight the need for managers to consider participative practices and fairness issues as they develop standards of performance for employees, especially those in lower ranking positions. Recognizing that employees at different ranks within the organization respond differently to various factors can lead to the development of more effective implementation plans for programs such as the BSC. The results of this study suggest LOWRANK employees are more concerned about the perceived fairness of the process surrounding the establishment of performance standards, while employees in higher ranking positions are more concerned with the feedback received about their performance.

This study is subject to the following limitations that should be considered when interpreting its results. First, the motivation construct used in the study is somewhat unique. Because of the aforementioned limitations set by bank officials in the wording of this construct, wording for the question was compromised. To the extent that "feedback" is a vague term that can have several different meanings (for example, verbal recognition, monetary rewards, and/or promotion consideration), we cannot draw conclusions as to the effectiveness of specific types of motivating rewards or what employees were thinking when answering this question. Yet, we contend that the question does capture the impact of the motivation associated with feedback related to performance on the BSC and leaves open for investigation the effect of different types of motivation in a budget-setting environment. Second, we used single-item measures for our attitudinal constructs. This again represents a compromise and in lieu of validated multi-item measures, we adopted questions very similar in wording to the measures used by Fisher et al. (2002). To the extent that our single-item measures do not properly capture employee attitudes, our results are to be interpreted with care. Next, we use attitudinal data collected during the seventh month of the BSC and performance data as of the end of the seventh month. Monthly updates on their scorecard performance are available, and therefore the possibility exists that employees were aware of their performance through 6 months on the BSC while taking the survey. This leaves open the possibility prior knowledge of performance-affected employees' perception of the process. Finally, the timeframe from the time the BSC was implemented until the time we collected survey data was less than a year. The possibility exists that the program is still too new to reflect permanent attitudes or performance data, and therefore further research is needed to investigate and better understand the impact of time as it relates to this study.

Aside from addressing the weaknesses associated with the limitations of the study, the findings from this study suggest several areas for future

research. While often difficult to achieve, replicating and extending experimental findings in the field is a valuable service to the academic community, especially if the ability to extend the research exists. In this study, we observed different results for fairness and motivation variables based on whether the employee was in a lower or higher ranking position within the organization. Future research can further explore the impact of ranking within the organization as it relates to the relationship between attitudes, behaviors, and performance. As discussed above, motivation and feedback can come in various forms, each with differing effects on attitudes and performance. Future research can explore the effect of different types of motivation on attitudes and performance. Finally, this study's findings have a bearing on the development of implementation procedures for budgets, including the establishment of BSCs for employees in the organization. Future research can continue to explore factors that result in improved practices for managers involved in the establishment of performance standards.

NOTES

1. Voice is defined as the ability of subordinates to be involved in a decision process by communicating their views to their superiors (Leventhal, 1980). A voice vote occurs when the subordinate is allowed to express to the superior whether the evaluation measures are acceptable. The subordinates' voice and vote in setting their standards give them a degree of perceived process control (Lindquist, 1995).

2. It is noted here that these expectations are similar to the expectations established and supported in Fisher et al. (2002). We review this literature and theory in order to support the forthcoming hypotheses.

3. See, for example, Kessler and Ashton (1981); Nelson (1993); Kluger and DeNisi (1996); Ashford and Cummings (1983); and Pritchard, Jones, Roth, Stuebing, and Ekeberg (1988).

4. While a total of 368 employees returned surveys, only 221 of the responses were from employees that were participants in the initial rollout of the BSC program (the program was not rolled out evenly across the organization). The 58 responses were removed in order to control for the possibility that performance standards were actually set at an unfair level. See the discussion in the *Performance Measurement* section for a more comprehensive discussion of these cases.

5. Responses were collected using a seven-point scale anchored by 1 = "Strongly Disagree" and 7 = "Strongly Agree".

6. Separating responses in this manner divided respondents into a group with no opinion or a disagreement with their target level (responses 1–4) and a group expressing at least a small degree of agreement with their performance target (5–7).

7. In an attempt to mitigate these concerns, we include only cases where respondents indicated their supervisors were aware of their performance capabilities. Further, our statistical analyses do not compare performances across high- and

low-ranking groups, only within each grouping. This lessens the opportunity for contextual differences in how performance is measured between high- and low-ranking to make a difference in our findings.

8. We consulted with bank management to confirm this segregation of employees represented a significant difference in hierarchical rank within the organization.

REFERENCES

Aquino, K., Grover, S. L., Bradfield, M., & Allen, D. G. (1999). The effects of negative affectivity, hierarchical status, and self-determination on workplace victimization. *Academy of Management Journal*, *42*, 260–272.

Ashford, S. J., & Cummings, L. L. (1983). Feedback as an individual resource: Personal strategies of creating information. *Organizational Behavior and Human Performance*, *32*, 370–398.

Brass, D. J., & Burkhardt, M. E. (1993). Potential power and power use: In investigation of structure and behavior. *Academy of Management Journal*, *36*, 441–470.

Brown, S. P. (1996). A meta-analysis and review of organizational research on job involvement. *Psychological Bulletin*, *120*, 235–255.

Chen, C. C. (1995). New trends in rewards allocation preferences: A Sino–US comparison. *Academy of Management Journal*, *38*, 408–428.

Chenhall, R. H., & Brownell, P. (1988). The effect of participative budgeting on job satisfaction and performance: Role ambiguity as an intervening variable. *Accounting, Organizations and Society*, *13*, 225–233.

Finkelstein, S., & Hambrick, D. C. (1990). Top management team tenure and organizational Outcomes: The moderating role of managerial discretion. *Administrative Science Quarterly*, *35*, 484–503.

Fisher, J., Frederickson, J. R., & Peffer, S. A. (2002). The effect of information asymmetry on negotiated budgets: An empirical investigation. *Accounting, Organizations and Society*, *27*, 27–43.

Hollenbeck, J. R., & Klein, H. J. (1987). Goal commitment and the goal-setting process: Problems, prospects, and proposals for future research. *Journal of Applied Research*, *72*, 212–220.

Homans, G. C. (1974). *Social behavior: Its elementary forms (revised edition)*. New York: Harcourt Brace Jovanovich.

Hunton, J. E., Hall, T. W., & Price, K. H. (1998). The value of voice in participative decision making. *Journal of Applied Psychology*, *83*, 788–797.

Kessler, L., & Ashton, R. H. (1981). Feedback and prediction achievement in financial analysis. *Journal of Accounting Research*, *19*, 146–162.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*, 254–284.

Kopelman, R. E. (1986). Objective feedback. In: E. A. Locke (Ed.), *Generalizing from laboratory to field settings*. Lexington, MA: Lexington Books.

Latham, G.P. (2004). The motivational benefits of goal-setting. *The Academy of Management Executive*, *18* (November 4), 126–129.

Lee, T. W., Locke, E. A., & Phan, S. H. (1997). Explaining the assigned goal-incentive interaction: The role of self-efficacy and personal goals. *Journal of Management*, *23*(4), 541–559.

Leventhal, G. S. (1980). What should be done with equity theory? In: K. J. Gergen, M. S. Greenberg & R. H. Willis (Eds), *Social exchange: Advances in theory and research* (pp. 27–55). New York: Plenum Press.

Lindquist, T. M. (1995). Fairness as an antecedent to participative budgeting: Examining the effects of distributive justice, procedural justice and referent cognitions on satisfaction and performance. *Journal of Management Accounting Research*, *7*, 122–147.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ: Prentice-Hall.

Martin, J., & Harder, J. W. (1994). Bread and roses: Justice and the distribution of financial and socioeconomic rewards in organizations. *Social Justice Research*, *7*, 241–264.

Mathieu, J. E., & Zajac, D. M. (1990). A review and a meta-analysis of the antecedents, correlates, and consequences of organizational commitment. *Journal of Applied Psychology*, *108*, 171–194.

Nelson, M. W. (1993). The effects of error frequency and accounting knowledge on error diagnosis in analytical review. *The Accounting Review*, *68*, 804–824.

Nouri, H., & Parker, R. J. (1996). The effect of organizational commitment on the relation between budgetary participation and budgetary slack. *Behavioral Research in Accounting*, *8*, 74–90.

Pritchard, R. D., Jones, S. D., Roth, P. L., Stuebing, K. K., & Ekeberg, S. E. (1988). Effects of group feedback, goal setting, and incentives on organizational productivity [Monograph]. *Journal of Applied Psychology*, *73*, 337–358.

Rousseau, D. M. (1995). *Psychological contracts in organizations: Understanding written and unwritten agreements*. Thousand Oaks, CA: Sage.

Schminke, M., Cropanzano, R., & Rupp, D. E. (2002). Organization structure and fairness perceptions: The moderating effects of organizational level. *Organizational Behavior and Human Decision Processes*, *89*, 881–905.

Vroom, V. (1964). *Work and motivation*. New York, NY: Wiley.

EXPECTANCY THEORY AS THE BASIS FOR ACTIVITY-BASED COSTING SYSTEMS IMPLEMENTATION BY MANAGERS

Ken C. Snead, Jr., Wayne A. Johnson and
Atieno A. Ndede-Amadi

ABSTRACT

Many studies, motivated by concerns for activity-based costing (ABC) implementation efforts being less than successful, have suggested that the lack of success in this area stems more from behavioral, as opposed to technical, factors. This concern for the behavioral aspects of systems implementation has also emerged from much of the more general information systems research examining determinants of implementation success. Accordingly, the purpose of this study is to determine if a popular process theory of motivation, expectancy theory, would be useful in explaining the motivation of managers to incorporate ABC information into their job. Data obtained from two experiments employing a judgment modeling methodology support the relevance of both the valence and force models of expectancy theory in this context. Further, the judgments provided by the subject managers suggest they perceive improved product cost accuracy as the most beneficial outcome of ABC use, followed by an

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 253–275

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14012-X

equivalent appreciation for both an enhanced ability to communicate the underlying economics of the firm and to identify non-value-added activities. Additionally, subject managers exhibited a greater concern for the possibility that obtaining the data to maintain the ABC system would be difficult and costly than they did for concerns that the ABC information would increase the level of complexity of the information that they use.

BACKGROUND

For well over a decade, the literature discussing activity-based costing (ABC) suggests that traditional cost allocation systems systematically distort product costs and consequently taint the information managers rely on for decision-making (Cooper & Kaplan, 1988; Lere & Colson, 2002). Accordingly, ABC costing procedures have been offered as a way to rectify this cost distortion dilemma, by providing more sophisticated costing information necessary for management to make more effective decisions in the areas of pricing, product mix, process improvements, etc. While the technical aspects of these costing procedures are intuitively appealing and are often responsive to the cost distortion dilemma, survey evidence suggests that many organizations adopting ABC are not experiencing the benefits anticipated (e.g., see Innes, Mitchell, & Sinclair, 2000). More specifically, some organizations report lack of success in implementing the new costing system (Innes & Mitchell, 1995), possibly representing the situation where the ABC system remains in the "analysis" stage, where the model is continually refined but never reaches the "action" stage, where the more sophisticated costing information is incorporated into the decision-making of management (Cooper, Kaplan, Maisel, Morrissey, & Oehm, 1992; Cokins, 2000). In essence, these concerns are the same as those that have motivated much of the era of information systems (IS) research investigating new IS implementation efforts (e.g., Swanson, 1988). Accordingly, it is important to realize that ABC constitutes a new IS and is thereby subject to the same user acceptance concerns that have typically plagued many new IS implementation efforts (Leonard-Barton, 1988).

This era of IS research has considered system utilization as one primary indicator of IS success (e.g., Barki & Huff, 1985; Ginzberg, 1981) and has consequently focused on the identification of variables associated with system use. These variables often include characteristics of the factors related to the IS implementation such as the user (e.g., personality variables, decision style)

(e.g., Chandrasekaran & Kirs, 1986), the system (e.g., accuracy, timeliness, relevance) (e.g., Baril & Huber, 1987), the organizational context (e.g., degree of centralization, size, growth rate) (e.g., Franz & Robey, 1986), and the implementation process (user involvement, top management support) (e.g., Fuerst & Cheney, 1982). Few consistent findings have emerged from this stream of research. While the variables mentioned above have been found to be associated with system use in one or more studies, their significant relationships with system usage have not been consistently demonstrated across studies. Accordingly, many of the post-hoc models differ in form and content, suggesting poor generalizability of results from this era of research (Nichols, 1981). This lack of consistency among results is attributed to the atheoretical approach of these studies (Ives & Olson, 1984).

Not surprisingly, the stream of ABC research motivated by the same user acceptance concerns appears to be focused on the same quest of identification of implementation factors associated with "successful" ABC implementation. Further, the factors considered in these studies are nearly identical to the user, system, organizational, and implementation process factors used in the prior IS era (see Table 1 in Anderson & Young, 1999 for a concise and comprehensive literature review summary). And of no further surprise, the resulting post-hoc models differ with respect to the set of variables included and path linkages deemed statistically significant.

An additional commonality between general IS research and research specific to the ABC context is that this gap between the ability to develop new managerial information technology, and to effectively use it is seen to result more from behavioral-related rather than system-related factors regardless of the specific work context (Lucas, 1975; Turner, 1982; Isaac, Zerbe, & Pitt, 2001; Anderson & Young, 1999; Chenhall, 2004). And while Chenhall (2004) notes, "… a difficulty exists in developing hypotheses as existing theories do not relate specific ABCM implementation factors to success, and empirical work varies in terms of effectiveness constructs, duration of implementation, and units of analysis," a common IS contention has been that the research in this area not only underutilizes existing knowledge in the behavioral sciences but also fails to tie implementation research to existing, more general models of work behavior (Robey, 1979; Ginzberg, 1980; Ives & Olson, 1984).

In response to the dearth of theoretically-based studies, Davis et al. (1989) employed a technology acceptance model to understand the system utilization behavior of individuals confronted with new technology. Their seminal findings indicate that system utilization is determined by individuals' behavioral intentions to use the system, which in turn are influenced by user

perceptions of the system's perceived usefulness and ease of use. Interestingly, these constructs are quite similar to the components of expectancy theory, a well established process theory of motivation, which has been found to generalize to a wide variety of contexts. Of specific interest to this study is the relevance of expectancy theory to the IS context, which was noted by researchers modeling the process of social change associated with operations research implementation (Vertinsky, Barth, & Mitchell, 1975). Additionally, Robey (1979) interpreted research results within the framework provided by expectancy theory in relating the extent of system use to user attitudes. Subsequently, other researches have noted the applicability of expectancy theory to IS-related research in a variety of work contexts. DeSanctis (1983) and Snead and Harrell (1994) found support for expectancy theory to model user acceptance of decision support systems, while Griffin and Harrell (1991) found expectancy theory was able to explain the motivation of a manager to adopt just-in-time management practices.

Accordingly, this study suggests that expectancy theory represents a reasonable conceptual framework for explaining the motivational force acting upon a manager to incorporate a newly developed ABC system into his/her job. This theory was chosen given its demonstrated generalizability to the IS and managerial contexts subsumed by ABC. An additional appeal of the theory is that it is a process theory of motivation and requires a within-person focus. These attributes are consistent with the findings of Anderson and Young (1999) who found significant respondent effects in their analysis and who allude to the value of process theories in this area of inquiry. And while most studies correlate use with financial performance measures (e.g., Ittner, Lanen, & Larcker, 2002), the focus of this study is to employ expectancy theory to identify the determinants of the intentions to use ABC; use is a necessary (but insufficient) condition that the ABC system will lead to operational, and ultimately, financial improvements.

EXPECTANCY THEORY

As originally formulated by Vroom (1964), expectancy theory explains how an individual chooses between alternative forms of behavior and continues to be used in a variety of contexts to provide a motivation theory-based explanation for individual behavioral intentions (Baker, Ravichandran, & Randall, 1989; Fusilier, Ganster, & Middlemist,1984; Harrell, Caldwell, & Doty, 1985; Nickerson & McClelland, 1989; Isaac et al., 2001). The essence of expectancy theory proposes that individuals will exert effort to do those

things that are expected to lead to outcomes that they value (or find attractive); it has often been referred to as an expectancy-value theory. Thus, an individual's perception of the likelihood that effort will lead to specific outcomes, coupled with the perception of the attractiveness of those outcomes, are proposed under the theory to be important determinants of that individual's motivation to perform particular acts. In the context of this study, the theory proposes that the motivation of a manager to use an ABC system is determined by his/her perception of the likelihood that the ABC system will lead to specific outcomes, coupled with the his/her perception of the attractiveness of those outcomes.

The original formulation of the theory incorporates the determinants of motivation into two models, the valence model and the force model (Vroom, 1964). The models are distinguished by the type of outcome each incorporates, as expectancy theory makes a distinction between two types of outcomes: first-level outcomes and second-level outcomes. A first-level outcome is the initial outcome directly expected from exerting effort, and in the context of this study, occurs when a manager incorporates the ABC information to a great extent in his/her job. This willingness to use the ABC information is one of the dimensions that DeLone and McLean (1992) suggest comprise the multi-dimensional construct of IS success. Second-level outcomes are outcomes or consequences to which the first-level outcome is expected to lead. In the setting of this study, a second-level outcome occurs as the result of a manager making extensive use of the ABC system in his/her job. Second-level outcomes are conceptually consistent with the individual impact (effect of information on user behavior) and organizational impact (effect of information on organizational performance) dimensions of IS success identified by DeLone and McLean (1992).

Whereas, the valence model describes how individuals consider the likelihood and attractiveness of second-level outcomes when forming assessments of attractiveness of a first-level outcome, the force model proposes that individuals form behavioral intentions by combining this attractiveness assessment with their perception of the likelihood that their effort will lead to the realization of the first-level outcome. Each model will be discussed in turn.

THE VALENCE MODEL

The valence model predicts the valence (attractiveness) of a first-level outcome, to an individual as a function of the sum of the products of the

valences of associated second-level outcomes and the strength of the perceived relationship between the first-level outcome and its associated second-level outcomes. This model therefore explains the valence of a particular act (or kind of behavior), as follows:

$$V_j = \sum_{k=1}^{m} V_k I_{jk} \tag{1}$$

where V_j is the valence to an individual of the first-level outcome j. V_k corresponds to the valence of the second-level outcome k, while I_{jk} is the strength of the relationship between the first-level outcome j and the second-level outcome k. V_j represents a manager's assessment of the attractiveness of using the ABC system extensively in performing his/her job, which is the attractiveness of the first-level outcome, j. V_k describes the attractiveness of each second-level outcome expected to result from the manager making maximum use of the ABC system, while I_{jk} describes the probability of each second-level outcome resulting from this maximum use.

Five second-level outcomes resulting from ABC system use are employed in this study, and are: (1) increased complexity of information used for decision-making; (2) more accurate identification of product costs; (3) increased ability to communicate underlying economics of the firm; (4) difficulty of obtaining needed information on a sustained basis; and (5) identification of "non-value-adding" activities. Support for each outcome is presented.

(1) *Increased complexity of information used for decision making* Typically, the single cost driver approach is considered simple, but not adequate. Consequently, other factors are introduced to more appropriately model the underlying economic complexity of the firm. While this is particularly important given today's complex manufacturing processes, the resulting complexity can overwhelm the user of the system (Cooper cited in Brinker, 1990; Keller & Krause, 1990). Accordingly, the potential user of an ABC system must consider this necessary increase in complexity when making an adoption decision.

(2) *More accurate identification of product costs* This benefit is identified as a key, desirable outcome related to the implementation of an ABC system, as there is perceived need for accounting systems capable of accurately capturing product cost information in the rapidly changing manufacturing environment (Howell & Soucy, 1988; Brunton, 1988; Anderson & Young, 1999; Cokins, 2000). This enables management to gain insight into the actual cost of producing particular products and servicing

specific customers (Cooper et al., 1992; Lere & Colson, 2002) in a variety of strategic decision contexts (make versus buy, product mix, outsourcing, etc.).

(3) *Increased ability to communicate the underlying economics of the firm* There has always been a problem understanding cost behavior patterns in a firm. In many cases, simplistic cost application models (such as all overhead applied on the basis of direct labor hours used) mislead users into thinking that application rates serve as a surrogate for the underlying economy (cost behavior) of the firm. ABC costing models, although more complex, do a much better job of communicating cost implications of business decisions. Also, the wide involvement of personnel in developing the ABC model provides some assurance that all concerned agree on the underlying economics of the firm (King, 1991). Thus, implementing an ABC system will improve communication (Cooper cited in Brinker, 1990), and will communicate cost information to all relevant groups, such as product design engineers (Turney cited in Brinker, 1990).

(4) *Difficulty of obtaining needed information on a sustained basis* By design, a more complex system is more costly (Turney cited in Brinker, 1990). In fact, case studies show that cost driver information is more costly to obtain (Cooper et al., 1992). Shillinglaw (1989) observes that management's unwillingness to adopt ABC concepts may be driven by reluctance to disrupt existing routines, the added cost of multiple driver systems, and the fact that management has not been convinced that added accuracy would produce significant incremental benefits. Often, these are not one-time costs. For example, distribution models must be maintained on an ongoing basis, and activity drivers must constantly be reviewed. As procedures change, new ABC models must be developed or cost accuracy will be lost. Accordingly, this ongoing maintenance can be a very costly process.

(5) *Identification of non-value adding activities* Process control is a major focus in the new manufacturing environment; this focus is congruent with ABC implementation requirements for the firm to identify activities that drive costs. Not only does this identify the cost of the activity, but it highlights the causes of work and helps identify improvement opportunities (Turney, 1992). Firms that implement ABC systems to identify product costs are frequently searching for ways to eliminate the need to perform some activities entirely (King, 1991). In fact, many ABC projects require not only an identification of product costs, but focus on process value analysis and cost reduction as well (Ostrenga, 1990). This

approach is consistent with modern management focus on the customer whereby what is important to the customer is what defines "value-added" as opposed "non-value-added" activities (Steimer, 1990; Cooper et al., 1992; Borthick & Roth, 1995; Carolfi, 1996).

THE FORCE MODEL

In the force model, Vroom (1964) hypothesizes that the motivational force acting on an individual to perform a particular act is a function of the sum of the products of the valences of the first-level outcome and the expectancy (probability) that the act will be followed by the attainment of this first-level outcome. The force model therefore explains the process by which an individual chooses to behave in a particular manner, as follows:

$$F_i = V_j E_{ij} \tag{2}$$

where F_i is the motivational force acting upon an individual to perform act i; E_{ij} is the expectancy or likelihood that act i will be followed by the first-level outcome j; V_j the valence of outcome j, is the link between Eqs. (1) and (2), as this variable is common to both equations. In the context of this study, an individual's decision to exert a particular effort level in order to use the ABC system to the maximum extent (F_i) indicates that individual's behavioral intentions, and reflects the level of motivation acting upon him/her. Moreover, that level of motivation to exert a particular effort level is explained by the attractiveness of using the ABC system (V_j) and the expectation that the choice of a particular effort level will result in being able to maximally use the ABC system (E_{ij}). A diagrammatic representation of the valence and force models is presented in Fig. 1, which highlights the V_j variable linkage of both models.

HYPOTHESES

The general research question underpinning this study is: Do the variables of the expectancy theory valence and force models explain the motivation of a manager to make voluntary use of an ABC system? The following hypotheses emerge from this question and from the discussion to this point:

H1a. The valence model will explain a manager's perception of the attractiveness of incorporating ABC information into his/her job.

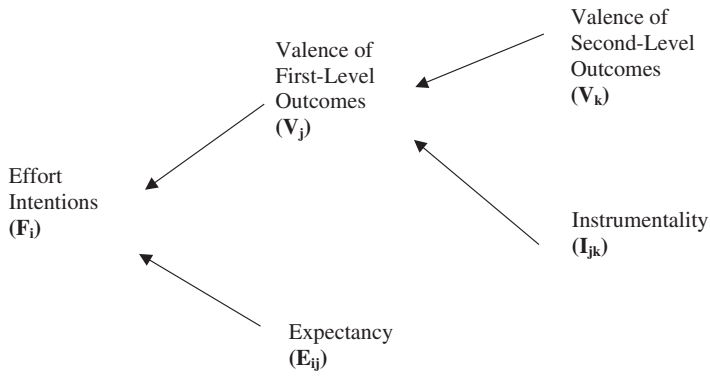


Fig. 1. Diagram of Expectancy Theory Variables.

H1b. On average, the second-level outcome valences will be positive for outcomes (2), (3), (5), and negative for outcomes (1), (4).

H1c. There will be no differences between the relative attractiveness for outcomes (2), (3), (5), and the relative unattractiveness for outcomes (1), (4).

H2. The force model will explain a manager's motivation to incorporate ABC information into his/her job.

The first three hypotheses (H1a–H1c) relate to the valence model of expectancy theory (Eq. (1)). H1a predicts that the attractiveness of using an ABC system will be determined by the attractiveness of the outcomes associated with its use and the strength of the manager's perception that these outcomes will result from using the ABC system. Presuming confirmation of H1a, H1b asserts that the sign of the second-level outcome valences (V_k) will correspond to the assertions in the literature that second-level outcomes involving more accurate identification of product costs (2), increased ability to communicate underlying economics of the firm (3), and identification of "non-value adding" activities (5) will be perceived as attractive (positively) by managers, while increased complexity of information used for decision-making (1), and difficulty of obtaining needed information on a sustained basis (4) will be perceived as unattractive (negatively) by managers. Presuming confirmation of H1b, H1c in null form merely invites an exploration of the relative measures of attractiveness and unattractiveness obtained for the second-level outcomes. H2 presumes support for H1a and is based upon the force model of expectancy theory (Eq. (2)). H2 predicts that a manager's motivation to use an ABC system will be determined by his/her perceptions

of the attractiveness of using the ABC system and perceptions of the probability that an effort to incorporate the ABC system into his/her job will be successful.

Support for H1a and H2 would imply that the variables of the valence and force models of expectancy theory are determinants of a manager's motivation to make voluntary use of an ABC system. As discussed earlier, this would imply that expectancy theory might provide an appropriate conceptual framework for identifying factors that determine a manager's intention to use an ABC system. This may, in turn, suggest practical approaches for increasing the voluntary utilization of ABC systems.

METHOD

Study Design

Vroom (1964) describes the force model as an individual choice model. Moreover, many researchers argue in favor of a within-person approach to studies that examine the theoretical relationships predicted by expectancy theory models (Harrell & Stahl, 1984; Kopelman, 1977; Wanous, Keon, & Latack, 1983; Wolf & Connolly, 1981). Studies that employ the within-person approach require measurements of effort level to be obtained from each participant under different expectancy-valence combinations. With this approach, the data for each individual are separately analyzed, usually by correlating the expectancy-valence motivation measures and effort level measures obtained for different situations. An advantage of the within-person approach is that many of the difficulties that can attenuate research findings, such as response bias, between-persons variance, and the failure to use ratio measurement scales are avoided (Arnold & Evans, 1979). It is noted that Anderson and Young (1999) determined that the individual should be the "unit of analysis" in this type of inquiry given their findings of significant respondent effects.

Noting the within-person nature of expectancy theory, Mitchell and Beach (1977) and Zedeck (1977) proposed that the judgment modeling approach frequently employed to examine cognitive issues (Ashton, 1982; Libby, 1981) represents a methodology congruent with the individual focus of expectancy theory. This methodological approach is now well established in expectancy theory research (Snead & Harrell, 1994; Baker et al., 1989; Butler & Womer, 1985; Harrell et al., 1985; Harrell & Stahl, 1984; Rynes &

Lawler, 1983; Stahl & Harrell, 1983) and was employed to gather the data for this study.

Judgment modeling involves providing an individual with a set of variables or cues with which to arrive at a particular judgment or decision. Multiple sets of these cues are presented, each representing a unique combination of strengths or values associated with the cues. A separate judgment is required from the individual for each unique combination of cue strengths presented. Various statistical techniques make use of the resulting multiple judgments and associated cue strengths in an effort to infer the strategy of cue usage employed by the individual in arriving at the judgments.

Measures

A judgment-modeling-based decision-making exercise was developed for the ABC system implementation for this study, which presented a number of situations, each representing a hypothetical, newly developed ABC system. Each situation required the participant to indicate the valence associated with making maximum use of the ABC system (V_j in Eq. (1)) and the level of effort the participant would exert to make maximum use of the ABC system (F_i in Eq. (2)). Maximum use is defined as relying upon the information generated by the ABC system to a great extent in performing the job and is consistent with the "extensive use of ABC" focus incorporated in the study by Ittner et al. (2002). The exercise instructions provided to the participants are shown in Exhibit 1, with a sample situation from the exercise presented in Exhibit 2.

The exercise was designed to incorporate each of the essential elements of the valence and force models (Eqs. (1) and (2)). The following discussion relates the elements of the sample situation provided in Exhibit 2 to these models. The five items presented prior to *Decision A* correspond to the second-level outcomes described in Vroom's (1964) valence model; the relevance of these five outcomes to ABC system implementation has already been indicated. *Decision A* represents the first-level outcome valence (V_j in Eqs. (1) and (2)), or the overall attractiveness of using a new ABC system to the maximum extent, given the likelihood that these five second-level outcomes would result from this use.

Further information is presented following *Decision A*. This likelihood information corresponds to E_{ij} in Eq. (2), and represents the expectancy that, if the individual exerts a great deal of effort, he/she will be able to incorporate the information generated by the ABC system maximally into his/her job.

Exhibit 1. Activity-Based Cost Accounting System Exercise Instructions.

Assume you are a product manager of a manufacturing company charged with the responsibility of deciding upon and recommending courses of action from among many alternatives. Your decisions and recommendations are based largely on your assessment of the cost impacts. A newly developed Activity Based Cost Accounting (ABC) system is available for your use which traces the cost of significant activities performed within your firm to products, customers, and other cost objectives. The system identifies the activities that incur costs (cost drivers), assigns a cost to each cost driver, and allocates these costs to specific cost objectives. Your use of this system is voluntary, and could range from minimum to maximum use. Minimum use essentially implies that you will continue to perform your job as you have always done, utilizing former cost allocation models. Maximum use means that you will rely upon the information generated by the ABC system to a great extent in performing your job.

Given this background, this exercise presents 32 situations; each different with respect to the likelihood of certain impacts associated with your making MAXIMUM use of the ABC system and with respect to the likelihood of your being able to incorporate the information generated by the ABC system to the MAXIMUM extent into your job. You are asked to make two decisions for each situation. You must first decide how attractive it would be for you to use the ABC system to the MAXIMUM extent (DECISION A). You must next decide how much effort you would exert to use the ABC system to the MAXIMUM extent (DECISION B). Use the information provided for each situation to reach your decisions. There are no "right" or "wrong" responses, so express your true beliefs openly. IT IS IMPORTANT THAT YOU PROVIDE RESPONSES FOR BOTH DECISION A AND DECISION B FOR ALL 32 SITUATIONS (situations are presented on both sides of the page); otherwise your responses will not be usable. Also, please provide the general information asked for on the last page of the exercise. Thank you for your participation in this project.

Next, participants arrive at *Decision B* by implicitly considering both the valence of maximum ABC system use (their *Decision A*) and the expectancy information. *Decision B* corresponds to F_i in the force model (Eq. (2)) and reflects the participant's motivation to use the new ABC system.

Two levels of likelihood (10% and 90%) were used for both the instrumentality associated with second-level outcomes (I_{jk}) and the expectancy variable (E_{ij}). The resulting situations, containing unique combinations of instrumentality and expectancy values, were developed by systematically varying the instrumentality and expectancy values from situation to situation. Furnishing each participant with multiple cases permits obtaining measures of motivation force levels under varied circumstances. This is a prerequisite for the within-person application of expectancy theory. The situations were presented in random order to avoid response bias.

Exhibit 2. Example Situation.

If you incorporate the information generated by the ABC system to the MAXIMUM extent into your job, the likelihood that–

the information that you use to make your decisions will be more complex is.....	LOW	(10%)
you will be able to more accurately identify your product costs is.....	HIGH	(90%)
you will be better able to communicate the underlying economics of the firm to subordinates and superiors is.....	HIGH	(90%)
it will be costly and difficult to obtain the needed information from employees on a continual basis is.....	HIGH	(90%)
you will be able to identify activities that do not add value to your products is.....	HIGH	(90%)

DECISION A: With the above outcomes and associated likelihood levels in mind, indicate the attractiveness to you of incorporating the information generated by the ABC system to the MAXIMUM extent into your job.

−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5
Very Unattractive										Very Attractive

FURTHER INFORMATION: If you exert a great deal of effort, the likelihood you will be able to incorporate the information generated by the ABC system to the MAXIMUM extent into your job is..... HIGH (90%)

DECISION B: Keeping in mind your attractiveness decision (DECISION A) and the FURTHER INFORMATION, indicate the level of effort you would exert to incorporate the information generated by the ABC system to the MAXIMUM extent into your job.

0	1	2	3	4	5	6	7	8	9	10
Zero Effort										Great Deal of Effort

The clear advantage of obtaining measures using this experimental approach is its ability to control for the many potential "nuisance effects" likely to become entangled with the behavioral factors impacting the motivation of managers to implement ABC. Specifically, differences among respondents with respect to uncertainties pertaining to the costs and benefits of the ABC and extent of system use, and differences with respect to the phase of adoption/system maturity are controlled for (Anderson & Young, 1999).

Subjects

Data to test the hypotheses were gathered from two experiments. The first experiment employed professional MBA students (PMBA) from two midwestern universities as subjects. The PMBA program is designed primarily for individuals who occupy full-time managerial positions. Each of these subjects completed the decision-making exercise (Exhibit 2), consisting of 32 situations, each situation representing a unique instrumentality/expectancy combination. A one-half fractional factorial design was incorporated into the five second-level outcomes shown prior to *Decision*

A ($2^5 = 32 \times 1/2 = 16$ combinations); these 16 combinations were then presented at two levels of expectancy (10 and 90%) to obtain 32 unique cases ($2^5 = 32 \times 1/2 = 16 \times 2 = 32$). A total of 54 individuals provided useable responses: 41 males and 13 females. The typical participant was 28 years of age, had been with his/her current employer for about 4 years, and supervised seven subordinates. Each had been exposed to the ABC subject matter as part of the graduate management accounting course.

The second experiment obtained data from the midwestern regional Institute of Management Accountants (IMA) members. A shorter version of the decision-making exercise used in the first experiment was mailed to 390 individuals on the regional IMA mailing list. This shorter version was identical in every respect to the instrument used in the first experiment, except that only 16 situations representing unique combinations of instrumentality/expectancy were used. For this group, a one-quarter fractional factorial design for the instrumentalities of the second-level outcomes was employed. The resulting eight situations ($2^5 \times 1/4$) were then presented at two levels of expectancy, generating the 16 situations. This shorter version permitted testing of the hypotheses and was thought to be more "inviting" for mailing list respondents. A total of 67 (17%) individuals responded, with 48 providing usable responses. The vast majority of unusable responses were decision-making exercises returned by individuals indicating they were retired, but still active with the IMA. Demographic information collected from respondents providing usable responses revealed that 12 were female and 34 male (two did not report), with the average age reported being 41 years. Respondents indicated they had been with their present employer for an average of 10 years and supervised six individuals.

RESULTS

Given the within-person methodological approach used, testing the two research hypotheses required a sequence of steps. These steps and corresponding results are discussed for each of the hypotheses.

H1a–H1c

H1a predicts that the valence model of expectancy theory will explain a manager's perception of the attractiveness (valence) of making maximum use of a new ABC system. This hypothesis was tested by estimating a multiple

regression model for each participant, as the individual is the appropriate unit of analysis due to the within-person nature of expectancy theory. *Decision A* (V_j in Eq. (1)) serves as the dependent variable, with the five second-level outcomes instrumentalities (I_{jk} in Eq. (1)) serving as the independent variables. Given the orthogonal nature of the research design, the resulting standardized regression coefficients (betas) represent the relative attractiveness of each of the corresponding second-level outcomes to each subject in arriving at *Decision A*. Thus, these betas represent second-level outcome valences, which are the V_k terms in Eq. (1) (Stahl & Harrell, 1983). Results are reported separately for each subject group.

The resulting valence regression model estimation procedure for the IMA group revealed that only one of the participants' models was not statistically significant (at the 0.05 level). Exhibit 3 contains the mean and median R^2 values for the remaining 47 participants having statistically significant models. As indicated, the average R^2 of the 47 significant regression models is 0.86. With respect to the PMBA group, all but five valence regression models achieved a significance level of 0.05 or less. Exhibit 3 indicates the average R^2 of the remaining 49 models is 0.71. Taken together, both experiments provide evidence of the explanatory power of Vroom's (1964) valence model in this ABC system implementation context. Accordingly, H1a is supported as the second-level outcome valences and their associated instrumentalities explain a manager's attractiveness assessment of using the ABC system.

Exhibit 3 presents information pertinent to the examination of H1b and H1c as it reports the mean and median beta (second-level outcome valence, V_k) for each of the five second-level outcomes, referenced as $V_1 - V_5$. The average participant found attractive those outcomes involving improved product cost accuracy, increased ability to communicate the underlying economics of the firm, and identifying non-value-added activities (V_2, V_3, and V_5, respectively). The remaining outcomes involving increased complexity of information and the costs associated with obtaining needed information were perceived as unattractive (V_1 and V_4, respectively). Recalling that these betas represent the relative attractiveness associated with each outcome, H1b is supported for both the PMBA and IMA groups.

Further, pairwise comparisons of the mean values for $V_1 - V_5$ were conducted to examine H1c. For those outcomes perceived as attractive, results revealed that V_2 is larger than either V_3 or V_5 ($p < 0.01$), and that there is no difference between V_3 and V_5 ($p > 0.07$). For those outcomes perceived as negative, $V_4 < V_1$ ($p < 0.01$). These results are consistent for both the PMBA and IMA groups and indicate that managers placed the highest measure of

Exhibit 3. Valence Model Regression Results.

	IMA GROUP (n = 47)		PMBA GROUP (n = 49)	
	Mean	Median	Mean	Median
R^2 (adj)	.86(.78)	.88(.82)	.71(.65)	.73(.68)
V_1	−.12	−.12	−.09	−.13
V_2	.58	.60	.50	.48
V_3	.31	.29	.32	.34
V_4	−.25	−.26	−.20	−.25
V_5	.34	.33	.37	.40

V_1- valence of increased information complexity
V_2- valence of improved product cost accuracy
V_3- valence of improved ability to communicate underlying economics
V_4- valence of difficulty of obtaining information
V_5- valence of identifying non-value-added activities

attractiveness on increased product cost accuracy, and were indifferent between the benefits associated with enhanced ability to communicate the firm's underlying economics and to identify non-value-added activities. In addition, managers viewed more negatively the potential for increased cost/difficulty associated with the continual information gathering requirements than they did the possibility of dealing with more complex information.

H2

H2 predicts that the participants' motivation to implement the ABC system will be a function of the product of the first-level outcome valence (V_j in Eqs. (1) and (2)) with expectancy (E_{ij} in Eq. (2)). The extent to which individuals employ this information multiplicatively (as indicated by Eq. (2)), as opposed to additively, is an ongoing issue in expectancy theory research. Accordingly, this issue was considered in conjunction with the examination of H2.

As with H1a, regression analysis was employed on a subject-by-subject basis to examine H2. The effort decision (*Decision B*) was treated as the dependent variable, with V_j, E_{ij}, and the $V_j{*}E_{ij}$ interaction comprising the independent variables. The significance level of the *t*-statistic associated with the interaction term was examined to determine if it offered significant incremental explanatory power over the additive combination of V_j and E_{ij}. Again, analysis was conducted separately for each subject group.

The results of the analysis performed for the IMA subjects indicate that only 36% of the subjects made use of the multiplicative combination of

expectancy and valence as implied by H2. Further, the average incremental increase in explanatory ability that resulted from including the interaction term was only about 0.04. The corresponding analysis for the PMBA group reveals similar results. Approximately 43% of the subjects in this group incorporated the multiplicative combination of valence/expectancy when forming their effort decisions. However, the average increment to explanatory power from the presence of this interaction term was only 0.03. Taken together, these findings support suggestions that the force model should be modified to reflect the fact that many individuals do not employ (or only marginally employ) multiplicative information-processing procedures when forming motivational level decisions (Stahl & Harrell, 1981).

Exhibit 4 presents the mean and median values for the R^2 and standardized regression coefficients resulting from the force model regression analysis when only the additive main effects are considered. Again, results are presented separately for each group. Mean R^2's for the IMA and PMBA groups are both 0.78. Further, the mean betas for the first-level outcome valence (b_1) for the IMA and PMBA groups are 0.75 and 0.77, respectively, and the mean betas for the expectancy term (b_2) for IMA and PBMA subjects are 0.32 and 0.30, respectively. Results of both experimental groups indicate the participants' motivation toward utilizing ABC information can be explained by the additive combination of expectancy and valence. Moreover, the mean beta information suggests that subjects were influenced more by their perceptions of the attractiveness of ABC system utilization (b_1) than by their expectations that effort would lead them to successfully incorporate the ABC information into their job (b_2). A statistical comparison of these betas reveal that $b_1 > b_2$ for both groups ($p < 0.01$).

These results indicate that when a within-person approach is used, both the valence and expectancy variables of the force model are significant determinants of an individual's motivation to use a new ABC system. However, the results do not support the multiplicative information processing implied by H2, as approximately only 40% of the participants employed multiplicative information-processing procedures in arriving at their motivational level decisions.

LIMITATIONS OF THE STUDY

The source of the strength of this study, its experimental nature, is also the primary source of its limitations. Limitations of this study include hindrances to generalizability, and involve the experimental nature of the task

Exhibit 4. Additive Force Model Regression Results.

	IMA GROUP (n = 47)		PMBA GROUP (n = 49)	
	Mean	Median	Mean	Median
R^2 (adj)	.78(.76)	.84(.82)	.78(.77)	.82(.80)
b_1	.75	.78	.77	.78
b_2	.32	.35	.30	.30

b_1- weight placed on valence/attractiveness of the ABC system
b_2- weight placed on the expectancy of maximum incorporation of ABC information into job

and the method of subject selection. Regarding the task, subjects' responses were gathered experimentally rather than from observation in a real world setting. Thus, only subjects' intentions were measured, and not their actual behavior. Further, the experimental task involved a limited number of outcomes; additional relevant outcomes that were not considered may also be influential. In addition, there is hindrance to external validity due to the non-random selection of subjects, as they (particularly the PMBA students) were selected based on their availability. The low response rate from the IMA membership mailing is attributed to misconceptions on the part of potential respondents as to the time required to complete the instrument. Due to the length of the instrument, it is feasible that responding to each of the 16 situations appeared "uninviting," and consequently discouraged individuals from starting the exercise. However, the corroborating results from the PMBA subject group attenuate non-response bias concerns for the IMA group.

DISCUSSION

The ongoing concerns for investments related to the analysis and development of ABC models to pay economic dividends will only be addressed when managers are motivated to incorporate and act upon ABC information. This same issue has been addressed more generally in the management IS success stream of literature, with some researchers having suggested that expectancy theory provides a theoretical framework capable of explaining the motivation of managers to adopt new IS. The results of the two experiments in this study suggest that the valence and force (additive form)

models of expectancy theory do a reasonable job of explaining the cognitive process a manager evokes when forming ABC implementation decisions.

Valence model results (H1a—H1c) generated relative beta sizes for second-level outcomes that suggest managers perceive improved product cost accuracy as most important, followed by an equivalent appreciation for enhanced ability to communicate underlying economics and to identify non-value-added activities. Additionally, managers exhibited greater concern for the possibility that the continual information demands of the ABC could be difficult and costly to obtain. Of lesser concern to managers is the likelihood that the ABC information would be more complex to process. This pattern of betas suggests that managers may be primarily influenced by outcomes they perceive to be more directly associated with profitability. Accordingly, they may see greater product cost accuracy as having more pervasive benefits in the decision areas of pricing, product mix, outsourcing, etc., while their larger concern for increased difficulty and costs associated with obtaining necessary information is likely perceived to have a more direct negative impact on profits than dealing with more complex information. Relatedly, the responses of those subjects receiving situations involving the highest likelihood of both benefits and costs (90% likelihood for all outcomes) had a mean attractiveness rating (*Decision A*) of $+2.52$ on the -5 to $+5$ scale, which was statistically greater than zero ($p < 0.01$). Thus, the attractiveness of the potential improvements in the areas of product cost accuracy, communication, and process improvements appear to outweigh concerns for increased complexity and the difficulty of acquiring ABC information. Accordingly, managers appear to be willing to absorb the negative aspects of ABC use in order to receive benefits.

Additive force model results (H2) generated relative beta sizes suggesting that managers forming effort to use decisions are more influenced by perceptions of the attractiveness of ABC system utilization than by their expectations that effort would lead to them successfully incorporate the ABC information into their job. Accordingly, this emphasis on attractiveness suggests managers need to be most convinced that their use of the information will likely lead to benefits associated with improved costing accuracy, etc., than they need convincing that they will be able to use the system.

The demonstrated relevance of expectancy theory for the ABC implementation context provides a theoretical framework for both the interpretation of prior research findings and for guiding future research. For example, the benefits of top management support, training, and user involvement on the implementation process may be in improving both the manager's expectancy of ability to use the system and perceptions that this

use is linked with desired outcomes. And based on the findings of this study, improving the belief that use will lead to desired outcomes (instrumentality) is likely to be most influential on the manager's behavioral intentions to use the ABC system.

Further, the perceived attractiveness of these outcomes related to system use (second-level outcome valences) are likely to vary with the type of position held by the manager. For example, production management may be most concerned with complexity and accuracy type outcomes; support functions may be focused more on identification of non-value-adding activities, with finance functions most interested in continual costs of maintaining the ABC system. The potential for these second-level outcome valences to vary may be the source of the cognitive conflict construct that Chenhall (2004) found to intervene between behavioral implementation factors and beneficial outcomes. Future research could employ expectancy theory as the underpinning to more directly test the above assertions. Additionally, the theory could provide the framework to examine the adoption behavior of managers as a function of the degree to which they perceive potential ABC outcomes are congruent with organizational strategies and reward structures.

REFERENCES

Anderson, S., & Young, S. (1999). The impact of contextual and process actors on the evaluation of activity-based costing systems. *Accounting, Organizations, and Society*, *24*, 525–559.

Arnold, H., & Evans, M. (1979). Testing multiplicative models does not require ratio scales. *Organizational Behavior and Human Performance*, *24*(1), 41–59.

Ashton, R.H. (1982). Human information processing in accounting. *Studies in accounting research #17*, Sarasota, FL: American Accounting Association.

Baker, D., Ravichandran, R., & Randall, D. (1989). Exploring contrasting formulations of expectancy theory. *Decision Sciences*, *20*(1), 1–13.

Baril, W., & Huber, G. (1987). Decision support systems for ill-structured problems. *Decision Sciences*, *18*(3), 350–372.

Barki, H., & Huff, S. (1985). Change, attitude to change, and decision support system success. *Information and Management*, *9*(5), 262–268.

Borthick, A., & Roth, H. (1995). Accounting for time: Reengineering business processes to improve responsiveness. In: M. Young (Ed.), *Readings in management accounting*. Englewood Cliffs, NJ: Prentice-Hall.

Brinker, B. J. (1990). *Emerging practices in cost management*. Boston, MA: Warren, Gorham and Lamont.

Brunton, N. M. (1988). Evaluation of overhead allocations. *Management Accounting*, *70*(July), 22–26.

Butler, J., & Womer, K. (1985). Hierarchical vs. non-nested tests for contrasting expectancy-valence models: Some effects of cognitive characteristics. *Multivariate Behavioral Research*, *20*(3), 335–352.

Carolfi, I. (1996). ABM can improve quality and control costs. *Cost and Management*, *70*(May), 12–16.

Chandrasekaran, G., & Kirs, P. (1986). Acceptance of management science recommendations: The role of cognitive styles and dogmatism. *Information and Management*, *10*(3), 141–147.

Chenhall, R. (2004). The role of cognitive and affective conflict in early implementation of activity-based cost management. *Behavioral Research in Accounting*, *16*, 19–44.

Cokins, G. (2000). Overcoming the obstacles to implementing activity-based costing. *Bank Accounting and Finance (Euromoney Publications PLC)*, *14*(1), 47–52.

Cooper, R., & Kaplan, R. S. (1988). How cost accounting distorts product costs. *Management Accounting*, *69*(April), 20–27.

Cooper, R., Kaplan, R., Maisel, L., Morrissey, E., & Oehm, R. (1992). *Implementing activity-based cost management: Moving from analysis to action*. Montvale, NJ: Institute of Management Accountants.

Davis, F., Bagozzi, R., & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003.

DeLone, W., & McLean, E. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95.

DeSanctis, G. (1983). Expectancy theory as an explanation of voluntary use of a decision support system. *Psychological Reports*, *52*(1), 247–260.

Franz, C., & Robey, D. (1986). Organizational context, user involvement, and the usefulness of information systems. *Decision Sciences*, *17*(3), 329–356.

Fuerst, W., & Cheney, P. (1982). Factors affecting the perceived utilization of computer-based decision support systems in the oil industry. *Decision Sciences*, *13*(4), 554–569.

Fusilier, M. R., Ganster, D. C., & Middlemist, R. D. (1984). A within-person test of the form of the expectancy theory model in a choice context. *Organizational Behavior and Human Performance*, *34*(3), 323–342.

Ginzberg, M. (1981). Early diagnosis of MIS implementation failure: Promising results and unanswered questions. *Management Science*, *5*(4), 459–478.

Ginzberg, M. J. (1980). An organizational contingencies view of accounting and information systems implementation. *Accounting, Organizations, and Society*, *27*(4), 369–382.

Griffin, L., & Harrell, A. (1991). An empirical examination of managers' motivation to implement just-in-time procedures. *Journal of Management Accounting Research*, *3*, 98–112.

Harrell, A., Caldwell, C., & Doty, E. (1985). Within-person expectancy theory predictions of accounting students' motivation to achieve academic success. *The Accounting Review*, *60*(4), 724–735.

Harrell, A., & Stahl, M. J. (1984). Modeling managers effort-level decisions for a within-persons examination of expectancy theory in a budget setting. *Decision Sciences*, *15*(1), 52–73.

Howell, R. A., & Soucy, S. R. (1988). Cost accounting in the new manufacturing environment. *Management Accounting*, *69*(August), 42–48.

Innes, J., & Mitchell, F. (1995). A survey of activity-based costing: A survey of CIMA members. *Management Accounting Research*, *6*(2), 137–153.

Innes, J., Mitchell, F., & Sinclair, D. (2000). Activity based costing in the UK's largest companies: A comparison of 1994 and 1999 survey results. *Management Accounting Research*, *11*(3), 349–362.

Isaac, R. G., Zerbe, W. J., & Pitt, D. C. (2001). Leadership and motivation: The effective application of expectancy theory. *Journal of Managerial Issues*, *13*(2), 212–227.

Ittner, C., Lanen, W., & Larcker, D. (2002). The association between activity-based costing and manufacturing performance. *Journal of Accounting Research*, *40*(3), 711–726.

Ives, B., & Olson, M. (1984). User involvement and MIS success: A review of research. *Management Science*, *30*(5), 586–603.

Keller, D. E., & Krause, P. (1990). World class down on the farm. *Management Accounting*, *71*(May), 39–45.

King, A. M. (1991). *The current status of activity-based costing: An interview with Robin Cooper and Robert S. Kaplan. Management Accounting*, *73*(September), 22–26.

Kopelman, R. (1977). Across-individual, within-individual and return on effort versions of expectancy theory. *Decision Sciences*, *8*(4), 651–662.

Leonard-Barton, D. (1988). Implementation characteristics of organizational innovation. *Communications Research*, *15*, 603–631.

Lere, J. C., & Colson, R. H. (2002). Selling activity-based osting. *CPA Journal*, *72*(3), 54–55.

Libby, R. (1981). *Accounting and human information processing: Theory and application*. Contemporary topics in accounting series. Englewood Cliffs, NJ: Prentice-Hall.

Lucas, H. C. (1975). *Why information systems fail*. New York: Columbia Press.

Mitchell, T., & Beach, L. (1977). Expectancy theory, decision theory, and occupational preference and choice. In: M. Kaplan & S. Schwartz (Eds), *Human judgement and decision processes in applied setting* (pp. 203–226). New York: Academic Press.

Nichols, M. (1981). A behavioral analysis for planning MIS implementation. *MIS Quarterly*, *5*(1), 57–65.

Nickerson, C., & McClelland, G. (1989). Across-persons versus within-persons test of expectancy-valence models. *Journal of Behavioral Decision Making*, *2*, 261–270.

Ostrenga, M. R. (1990). Activities: The focal point of total cost management. *Management Accounting*, *71*(February), 42–49.

Robey, D. (1979). User attitudes and management information system use. *Academy of Management Journal*, *22*(3), 527–538.

Rynes, S., & Lawler, J. (1983). A policy capturing investigation of the role of expectancies in decision to pursue job alternatives. *Journal of Applied Psychology*, *68*(4), 620–631.

Shillinglaw, G. (1989). Managerial cost accounting: Present and future. *Journal of Management Accounting Research*, *1*, 33–46.

Snead, K., & Harrell, A. (1994). An application of expectancy theory to explain a manager's intention to use a decision support system. *Decision Sciences*, *25*(4), 499–513.

Stahl, M. J., & Harrell, A. M. (1981). Modeling effort decisions with behavioral decision theory: Toward an individual differences model of expectancy theory. *Organizational Behavior and Human Performance*, *27*(3), 303–325.

Stahl, M., & Harrell, A. (1983). Using decision modeling to measure second-level valences in expectancy theory. *Organizational Behavior and Human Performance*, *32*(1), 23–34.

Steimer, T. E. (1990). Activity-based accounting for total quality. *Management Accounting*, *72*(October), 39–42.

Swanson, E. (1988). *Information system implementation: Bridging the gap between design and utilization*. Homewood, IL: Irwin.

Turner, J. A. (1982). Observations on the use of behavioral models in information systems research and practice. *Information and Management*, *5*(4&5), 207–213.

Turney, P. B. (1992). Activity-based management. *Management Accounting*, *73*(January), 20–25.

Vertinsky, I., Barth, R. T., & Mitchell, V. F. (1975). A study of OR/MS implementation as a social change process. In: R. L. Schultz & D. P. Slevin (Eds), *Implementing operations research/management science* (pp. 253–272). New York: American Elsevier.

Vroom, V. C. (1964). *Work and motivation*. New York: Wiley.

Wanous, J. P., Keon, T. L., & Latack, J. C. (1983). Expectancy theory and occupational/organizational choices: A review and test. *Organizational Behavior and Human Performance*, *32*(1), 66–86.

Wolf, G., & Connolly, T. (1981). Between-subject designs in testing expectancy models: A methodological note. *Decision Sciences*, *12*(1), 39–45.

Zedeck, S. (1977). An information processing model and approach to the study of motivation. *Organizational Behavior and Human Performance*, *18*(1), 47–77.

This page intentionally left blank

DYSFUNCTIONALITY IN PERFORMANCE MEASUREMENT WHEN OUTPUTS ARE DIFFICULT TO MEASURE: A RESEARCH NOTE

Robert Greenberg and Thomas R. Nunamaker

ABSTRACT

Issues of performance measurement are ubiquitous in modern organizations and are often concerned with evaluations of outputs or efficiency (which encompasses both inputs and outputs) of an entity or process. Examples of output measures include revenue generated, defective units produced, on-time shipments, etc. Efficiency examples include standard cost variances, machine up-time rate, and efficiency scores from input–output models such as Data Envelopment Analysis (DEA).

Difficult-to-measure outputs are often included even though they cannot be measured with precision. When outputs of a production process are not easy to measure, serious dysfunctional decision-making can be expected and these problems may be particularly acute when efficiency measurements from input–output models are directly tied to rewards and incentives. Both for-profit firms and public sector organizations may share output measurability problems.

In this paper, we examine the possible problems of using input–output models (such as DEA) when outputs are difficult to quantify within an

Advances in Management Accounting
Advances in Management Accounting, Volume 14, 277–289

ISSN: 1474-7871/doi:10.1016/S1474-7871(05)14013-1

agency theory perspective and illustrate the potential problems using recent proposals in the UK for evaluating and rewarding police unit performance. We conclude that although input–output models, particularly those such as DEA may be useful as a diagnostic tool to assist decision-makers in altering future operating strategies and policies, it has serious limitations when rewards and incentives are attached to the DEA performance evaluations. In our view, overreliance on mechanical, formula-based approaches is potentially a serious threat to improving performance in these situations.

INTRODUCTION

In this paper we discuss concepts of performance measurement and how they relate to input–output models used for rewards and incentives. Of particular importance is the issue of data manipulation and its potential impact on efficiency measures when output measurability is low. Problems of low output measurability occur in functions such as research and development (R&D) (Abernathy & Brownell, 1987) in for-profit organizations and are predominant in public sector organizations.

In this paper we use an agency theory perspective to examine performance measurement and resulting resource allocation decisions using DEA, an input–output model that has been widely applied in both for-profit and not-for-profit organizations (Emrouznejad, 1995–2001).[1] Our central observation is that although DEA may be useful as a diagnostic tool, e.g., assisting decision-makers in altering future strategies for conducting operations, it has serious limitations when economic consequences are attached to the DEA performance evaluations. That is, serious dysfunctional decision-making can be expected. In our view, overreliance on mechanical, formula-based approaches is potentially a serious threat to improving organizational performance.

For our discussion, we use as a backdrop the recent proposal in the UK for implementing an incentive-based DEA performance evaluation system throughout police units in England and Wales. This report is termed the Public Services Productivity Panel (PSPP) (2000) report and selected, relevant elements are summarized in Stone (2002) and Drake and Simper (2003). Readers are referred to these publications for details of the PSPP report as they apply to performance measurement and resource allocation.

For purposes of our analysis, the readers need only to know that the DEA input–output model has been proposed in the UK on a national basis for evaluating *and rewarding* performance of police departments where output measurability is problematic. Similarly, where output measurability is low or at least problematic for certain for-profit activities such as R&D, advertising, maintenance, and product support, etc., our conclusions should continue to hold: rigid adherence to scalar efficiency measures as a basis for rewards can lead to unintended, unwanted behavior.

Observability is a key concept that affects the use of performance measures in two ways. First, because the underlying phenomena of interest are usually unobservable, surrogates must be found to proxy for the phenomena. Second, the unobservability of inputs affects the choice of control systems thereby affecting the choice of organizational form. These issues are discussed in order. We next address psychological responses by evaluators to the complexity and ambiguity inherent in performance evaluation. In the final section, we conclude with a discussion of the potential dysfunctional uses of input–output performance evaluation with particular reference to DEA.

CHARACTERISTICS OF MEASUREMENTS: PRINCIPALS AND SURROGATES

Performance measurements, like other measurements, are used to convey information about entities, objects, or phenomena. In discussing measurement, Ijiri (1975) distinguishes between *principals* and *surrogates*. He calls the underlying entity or phenomenon of interest the *principal*.[2] However, the usual case is that the principal is not directly observable and measurements are *surrogates* used to convey information about the principal.

Consider the following example in the context of policing. The input to the process includes the sacrifice of resources used to produce the policing outputs including resources such as the buildings used to house the police station, the vehicles (and their fuel, maintenance, etc.) used in policing, supervisors' time, office supplies, officers' time, etc. The actual sacrifice of these resources is not directly observable but a proxy (or surrogate) for the sacrifice is the *measured* cost reported by the accounting system.

Apart from the fact that the accounting system may fail to report a good measurement of the sacrificed resources, managers may manipulate the cost reported by the system by choosing *whether* to incur a cost, *when* to incur the

cost, or by outright *deception*. Deception may take many forms including failing to report a cost, reporting the cost in the wrong time period, or attributing the cost to an entity other than the one actually consuming the resources.

The important point is that because the principal is often unobservable, a surrogate is used to convey information about the principal. Measurement systems may fail when there is a low degree of correspondence between the principal and surrogate,[3] or when managerial incentives lead to dysfunctional management decision-making and/or deception.

The Control Model

Our view of a functional control model is similar to that outlined in the PSPP Report (p. 11) but explicitly recognizes the potential effect of exogenous factors. Fig. 1 depicts the process whereby resources are converted into performance of objectives. Resources are the *inputs* to the process that converts them into *outputs*. In the context of policing, the inputs might be labor hours, gasoline, and vehicle maintenance and the outputs might be hours of patrol. *Technical efficiency* refers to the quantity of output that is generated for a given level of input; higher technical efficiency is preferred. Measures of technical efficiency may be thought of as measuring the efficiency of the conversion of inputs into outputs as illustrated in Fig. 1.

A process is in place to meet one or more *performance objectives. Outcomes* of the process are differentiated from its *outputs*, in that outputs from

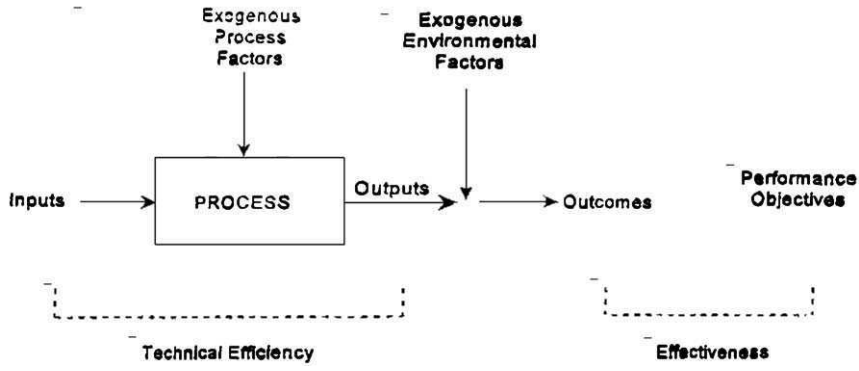


Fig. 1. Functional Control Model.

a process may meet, or fail to meet, the performance objectives for the process. Whether or not the performance objectives are met determines the *effectiveness* of the process as illustrated in Fig. 1. In the context of policing, the performance objectives might include limiting crime levels to specified targets and attainment of specified levels of community policing approval. A department might be technically efficient in producing patrol hours but might not be effective in meeting its performance objectives.

Fig. 1 also portrays the possible effects of factors outside the control of those managing the process. *Exogenous process factors* are those factors affecting the conversion of inputs into outputs. For example, the employment of relatively inexperienced officers requiring more supervision (necessitated by the retirement of experienced officers) might lead to unavoidable technical inefficiency that is outside the control of managers. *Exogenous environmental factors* include environmental factors that moderate the effectiveness of *outputs* in producing *outcomes* that meet *performance objectives*. For example, an increase in unemployment or social unrest may lead to an increased crime rate for a given level of output produced by a technically efficient process.

The level of resources that are inputs to the process may also limit the control model's effectiveness. Even though a process is technically efficient, the level of inputs may be insufficient to produce enough output to meet performance objectives. Moreover, resource inputs are often fixed in the short run by budgetary procedures and those in control of the process may be limited in their ability to respond to changes in exogenous environmental factors in the short run.

Fig. 1 simply portrays the process inputs and outputs as principals and does not explicitly recognize that surrogates must be found to provide their measurement. This has important consequences for any performance analysis system (such as DEA) that relies on the surrogates as the raw material for their analyses.[4]

Diagnostic and Accountability Uses of the Control Model

The control model may be usefully applied at the single department level by a manager who has no reporting requirements outside of the department. In this case, a prudent manager may use measures of technical efficiency to *monitor performance* and provide *diagnostic* information to improve performance. Comparison of outcomes to performance objectives will indicate if the department is effective. Information concerning technical efficiency

will be useful when considering changes in strategy, and strategy changes will be reflected in revised performance objectives. Effectiveness measurements will indicate the degree of success in implementing new strategies.

In the scenario just described, the manager has no reporting responsibility outside of the department. The more usual case, however, is that the manager is accountable to superiors for the prudent use of resources and the achievement of performance objectives. The measures of technical efficiency and effectiveness described above are often the measures reported for superiors for the purposes of *performance evaluation*. That is, rather than being used only by the manager to monitor and improve performance at the departmental level, these measures are being used by superiors to determine the quality of the manager's performance in directing the unit's scarce resources to achieve the unit's objectives.

When these measures are used either implicitly or explicitly as in the case of the UK proposal, as a basis for financial incentives (i.e., financial rewards and punishments), the measures become the basis for compensation contracting between the manager and superior. Because the subordinate manager's compensation and/or funding depends on these measures, there exists a conflict of interest between the subordinate and superior. The implications of this conflict have been discussed at length in accounting research conducted under the agency theory paradigm and as *accountability* by Ijiri (1975).

AGENCY THEORY AND PERFORMANCE EVALUATION

An intuitive summary and explanation of agency theory concepts is provided by Thornton (1984, 1985). In his papers, Thornton explains the difference in contracting options for team workers when inputs (e.g., employees' effort levels) are readily observable versus when such input levels are not observable.

When inputs are observable, agency theory predicts that for-profit corporations will be the preferred form of organization as owners will effectively monitor (or establish systems to monitor) the employees' efforts and prevent shirking behavior. The owners will establish such monitoring because they are the "residual claimants" of any profit generated by the firm.

What happens when inputs are not observable? Agency theory predicts that the optimal organizational form will not be the corporation, instead it

will be one in which mutual trust and stringent criteria for employment will be paramount. Thornton uses the example of accounting firms, which are almost universally organized as partnerships and not corporations. In this case, Thornton suggests the control system takes the form of careful selection of new employees and screening of candidates for advancement. If the right employees are hired, desired outcomes will follow.

A classic example of a unit with unobservable *inputs and outputs* is the R&D department. Indeed, Thornton's view reflects precisely the management control model followed by the R&D divisions of for-profit corporations (Abernathy & Brownell, 1987). Other corporate activities where *output* measurement is troublesome, although less acute than in R&D, include advertising, product support, and service.

In the case of the UK police departments (as with R&D departments in corporations), *inputs* and *outputs* are again not observable. How does one know if a police officer provides that extra attention to follow all leads on a crime report, which results in making witness contacts that assist in future police investigations? Yes, the output of such attention is unmeasurable but observability of the input (providing the extra attention on the initial case) is also problematic. When a police officer spends extra time consoling a crime victim, the output is not measurable nor can the input be observed and *monitored* in any practical way. Thornton makes the point that simple monitoring by individuals that are not owners of the firm leads to a logical paradox: the monitors need to be monitored, who also need to be monitored, and so on. Linkages of agency theory to matters of "data hardness" are examined next.

Accountability, Performance Measurement, and Hardness

Ijiri (1975) denotes the person or entity reporting performance measurements as the *accountor* and the person or group receiving the measures as the *accountee*. Accountability arises from the accountor's stewardship responsibility to the accountee and as Ijiri notes, "it is rather uncommon to have a situation where the interest of the accountor completely coincides with the interest of the accountee." The stewardship relationship (accountability) gives rise to performance measurement and, according to Ijiri, it is "impossible to discuss performance measurement without understanding the pressures that may be exerted by the entity (accountor) and the recipients of the measures (accountee) because of their self-interest."

These competitive interests necessitate performance measures that are well specified and verifiable to withstand pressures by the accountee and accountor to bias or dispute the measures.[5] Ijiri calls such measures hard.[6] Measures that may seem on their face to adequately track achievement toward performance objectives may be unsatisfactory because they result in "abusive use of performance measures" (Ijiri, p. 35) and conflict if they lack hardness.

Thus, when performance measures are reported for *accountability* purposes rather than for only *diagnostic* purposes, they must not only correspond to the principal for which they proxy, they must be well specified and verifiable in order to avoid conflict and attempts to bias. Within the incentive contracting literature, formal analytical results coincide with these intuitive observations (see, for example, Burgess & Metcalf, 1999).

Although the use of performance measures in evaluation is problematic, they are ubiquitous and there is apparently an irresistible compulsion to employ them. To quote Baker, Gibbons, and Murphy (1994, p. 1125), "Business history is littered with firms that got what they paid for." This issue is examined in the next subsection.

Ambiguity in Performance Evaluation: The Search for a Single Number

One needs to look no further than the current business periodicals and media to see that (past, current, and forecasted) profitability measures (e.g., net income, EPS, etc.) are of paramount importance when attempting to value a firm's stock. That is, when faced with a complex evaluative task, decision-makers tend to look for ways to simplify the problem. They search a summary measure (or relatively few measures) to simplify the task. This behavior has been the focus of research in psychology (Simon, 1955, 1956; Tversky & Kahneman, 1974): When confronted with complex, ambiguous decision tasks, individuals will rely upon various heuristics and rules of thumb to simplify the decision process. It appears that sometimes these rules of thumb are as good as more formal linear decision models (Todd, 2000); sometimes they are not (Harvey, 1998).

As might be expected, evaluation of an entity with multiple, difficult-to-measure inputs and outputs is a task of daunting complexity that motivates a decision-maker to seek simplification through use of a heuristic. The presence of multiple measures motivates the search for a summary measure to make the problem manageable.

The propensity to drift toward a single performance measure is well expressed by Hibbert who, in discussing Cox et al. (1992) (concerning the search for a single output measure in the healthcare sector), noted that "Those seeking to describe complex phenomena, or to take decisions based on them, will inevitably be drawn toward summary measures with an apparent scientific basis… ."

In the case of the UK police departments. Stone (2002, p. 11) observes that decision-makers in the UK are "currently looking for some single measure of efficiency to help in a revision of the present police funding formula…," rather than relying upon multiple indicators of various performance dimensions. The UK proposal is to use the efficiency score from the DEA model to evaluate and reward police department performance. One problem with the DEA efficiency score is that it is incomplete and subject to manipulation, which if anchored upon by evaluators can potentially result in dysfunctional decision-making (Nunamaker, 1985, 1988; Stone, 2002). The use of the DEA score is an example of the "Take the Best" heuristic discussed by Todd (2000).

Using this heuristic, the decision-maker searches for cues in the order of their perceived correlation with the decision criteria, and then selects the one with the highest perceived correlation for evaluative purposes. This approach is essentially a non-linear, non-compensatory decision rule; other cues have no impact on the decision. Importantly though, the successful use of simple decision heuristics (focusing on single efficiency scores from an input–output model such as DEA) depends critically on how well the decision cue (e.g., a DEA efficiency score) matches with the characteristics of the decision environment (Todd, 2000).

POTENTIAL DYSFUNCTIONALITIES OF INPUT–OUTPUT PERFORMANCE MEASUREMENT

Drawing together our prior arguments adds the notion that a single measure such as the DEA efficiency score may be used for resource allocation, as essentially proposed in the PSPP report. That is, incentives will be directly attached to the calculated performance rating. What is the logical and rational approach to performance reporting by the agent? The principal–agent model tells us that agents will maximize their own utility and suggests that a variety of techniques under the control of the agent will be used to manipulate the all-important performance score. Indeed, analytical work by

Baker (1992) indicates that the size of a piece-rate incentive and the efficiency of the payment contract depend upon the statistical relationship between the performance measure and the principal's objective. In an environment of information asymmetry (where the agent has more information than the principal) good, hard performance measures that mirror the principal's objectives are needed to avoid gaming behavior intended to bias the measure in favor of the agent.

The accounting research literature is replete with studies documenting income smoothing, earnings management, etc. that seek to "window-dress" the bottom line income number in the for-profit corporate setting.[7] When DEA or other summary evaluation technique is used as a control model where inputs and outputs lack hardness and reliability, why would we expect any less attempt at manipulative behavior?

Moreover, consider our observations in the context of opportunities for data manipulation. In particular, DEA reflects a strong non-compensatory, disjunctive decision model. In the words of Hogarth (1987, p. 76), with the disjunctive model, "A decision maker…will permit a low score on a dimension provided there is a very high score on one of the other dimensions. In other words, the candidate would be evaluated according to his or her best attributes regardless of the levels on the other attributes." This is exactly the rationale behind Pareto Efficiency used in the DEA model (Nunamaker, 1985). In the case of the UK scenario, a police department may be judged "efficient" because it scores well on one performance measure while scoring poorly on all others. Moreover, it may score well on the performance measure because of the difficulty of measurement or the choice of the particular performance measures included in the evaluations.

Difficulty of measurement may manifest because of (A) incongruence between the "principal and surrogate" performance measures, (B) measures may lack "hardness" allowing data manipulation, and (C) errors. Moreover, if unimportant input or output measures are included in the model, or important input or output measures are omitted, a police department might score well on only one measure and erroneously deemed "efficient" (and vice versa). These problems all contribute to a strong potential for erroneous evaluation and dysfunctional resource allocation decisions associated with rewards.

Measurability problems imply that mechanistic input–output control models cannot be automatically applied to activities lacking well-specified outputs. Control models strongly tied to competitive rewards suggest that those activities that can be reliably measured and compared are the important functions of the firm. By implication, any activity that cannot be

quantified and measured must be less important, and only those measured activities will be used for incentive contracting. Such an approach may work well when outputs are well specified and there is a high degree of correlation between the principal and surrogate measures.

However, when output measurability problems exist, the underlying control model should be altered to reflect these difficulties. This is particularly evident when control models are the driving force behind a system of rewards and incentives. Of note, the problem is not confined to simple random measurement error or statistical noise. Attempts to adapt DEA to a stochastic environment can be found in Ruggiero (2004) and Banker et al. (2004). The difficulty we identify in this paper employs assumptions used by Baker (1992) and Holstrom and Milgrom (1991) in their analytical studies. That is, given asymmetrical information problems and difficulties in observing an agent's behavior, there are strong incentives for agents to bias performance measures in their favor resulting in undesirable outcomes for the principal.

As an alternative to strict, mechanistic approaches, Ouchi's (1979) theory of management control focuses on "people" or "social" controls when inputs and outputs are not readily observable. The essence of this control strategy is summarized well by Eisenhardt (1985, p. 135). Compared to input–output performance analysis, the social control strategy asserts that when the task is vaguely defined and outcome measurability is low, "…control can be achieved by minimizing the divergence of preferences among organizational members. That is, members cooperate in the achievement of organizational goals because members understand and have internalized these goals. This strategy emphasizes people policies such as selection, training, and socialization." In a complementary fashion, Baker et al. (1994) demonstrate analytically that formula-based performance and incentive contracts are improved when subjective measures are incorporated into the reward scheme. Organizations and their constituents would be well served by greater consideration of such "people" control strategies for evaluation and resource allocation decisions, rather than focusing so heavily on deterministic models in hopes of simplifying inherently complex decision processes. At the very least, firms should consider incorporating greater subjective assessments in instituting performance measurement/reward systems.

NOTES

1. This represents a web site at www.deazone.com, which contains an extensive bibliography of published studies on all aspects of DEA. One subsection contains

references to DEA applications in corporate settings such as manufacturing, banking, and investment activities. Numerous public sector applications are also cited.

2. Confusion with the term "principal" as used in agency theory should be avoided. In agency theory, the principal is the residual claimant in the relationship where an agent is hired to work for the principal. In measurement theory, Ijiri uses the term principal to denote the often-unobservable phenomenon of interest for which a measurement is desired.

3. Obviously, the unobservability of the principal complicates the assessment of the degree of correspondence. The relative quality of surrogates are often debated on "logical" grounds although managers' incentives may color their arguments.

4. In the language of information processing, *GIGO* (garbage in, garbage out).

5. The accountor has incentive to bias performance reporting upward because of the adverse consequences associated with failure to meet performance objectives. The accountee has incentive to bias the performance measures downward because of perceived adverse effects to the accountee's reputation and political consequences associated with failing to meet performance objectives.

6. Hardness and objectivity, although related, are differing concepts. Objectivity refers to consensus among neutral observers whereas hardness refers to a measure's ability to resist competitive pressures by non-neutral observers to bias it upward or downward. See Ijiri (1975, Chapter 3), for an excellent discussion of objectivity, hardness, and their relationship.

7. Unless you have been living in a cave on some distant planet, it should be well known that such behavior is commonplace. Current empirical evidence of income manipulation to seek private gain is overwhelming in the public media with cases such as Enron, WorldCom, Tyco, etc.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Mervyn Stone and participants at the July 2000 meeting of the Official Statistical Section of the Royal Statistical Society, London, England, for their helpful comments. Additional appreciated comments were provided by participants at the AIMA conference, Monterey, CA, May 2003.

REFERENCES

Abernathy, M., & Brownell, P. (1987). Management control systems in research and development organizations: The role of accounting, behavior and personnel controls. *Accounting, Organizations and Society*, *22*(3/4), 233–248.

Baker, G. (1992). Incentive contracts and performance measurement. *The Journal of Political Economy*, *100*(3), 598–614.

Baker, G., Gibbons, R., & Murphy, K. (1994). Subjective performance measurements in optimal incentive contracts. *The Quarterly Journal of Economics*, *4*(November), 1125–1156.

Banker, R., Chang, H., & Cooper, W. (2004). A simulation study of DEA and parametric frontier models in the presence of heteroscedacity. *European Journal of Operational Research*, *153*(3), 624–640.

Burgess, S., & Metcalf, P. (1999). *Incentives in organizations: A selective overview of the literature with application to the public sector*. CMPO Working Paper series No. 00/16. Bristol, UK: University of Bristol.

Cox, D., Fitzpatrick, R., Fletcher, A., Gore, V., Spiegelhalter, D., & Jones, D. (1992). Quality-of-life assessment: Can we keep it simple (with discussion)? *Journal of the Royal Statistical Society, Series A*, *155*, 353–393.

Drake, L., & Simper, R. (2003). The measurement of English and Welsh police force efficiency: A comparison of distance function models. *European Journal of Operational Research*, *147*(1), 165–186.

Eisenhardt, K. (1985). Control: Organizational and economic approaches. *Management Science*, *31*(February), 134–149.

Emrouznejad, A. (1995–2001). *Ali Emrouznejad's DEA homepage*. Coventry, UK: Warwick Business School.

Harvey, J. (1998). Heuristic judgment theory. *Journal of Economic Issues*, *32*(1), 47–64.

Hogarth, R. (1987). *Judgement and choice*. New York: Wiley.

Holstrom, B., & Milgrom, O. (1991). Multi-task principal-agent analysis: Linear tasks, asset ownership and job design. *Journal of Law, Economics and Organizations*, *7*, 24–52.

Ijiri, Y. (1975). *Theory of accounting measurement*. Sarasota, FL: American Accounting Association.

Nunamaker, T. (1985). Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation. *Managerial and Decision Economics*, *6*(1), 50–58.

Nunamaker, T. (1988). Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation – reply. *Managerial and Decision Economics*, *9*, 256.

Ouchi, W. (1979). A conceptual framework for the design of organizational control mechanisms. *Management Science*, *25*(September), 833–848.

Public Services Productivity Panel (PSPP). (2000). *Improving police performance*. London, UK: HM Treasury.

Ruggiero, J. (2004). Data envelopment analysis with stochastic data. *Journal of the Operational Research Society*, *55*, 1008–1012.

Simon, H. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118.

Simon, H. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 99–118.

Stone, M. (2002). How not to measure the efficiency of public services (and how one might). *Journal of the Royal Statistical Society, Series A*, *165*(3), 1–17.

Thornton, D. (1984). A look at agency theory for the novice – Part 1. *CA Magazine*, *117*(II), 90–98.

Thornton, D. (1985). A look at agency theory for the novice – Part 2. *CA Magazine*, *118*(I), 93–100.

Todd, P. (2000). The ecological rationality of mechanisms evolved to make up minds. *American Behavioral Scientist*, *43*(6), 940–956.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(September), 1124–1131.

This page intentionally left blank