

THESE

1 2008



This is to certify that the dissertation entitled

SUBJECTIVE ADJUSTMENTS TO OBJECTIVE PERFORMANCE MEASURES: AN EMPIRICAL EXAMINATION IN COMPLEX WORK SETTINGS

presented by

i

i

ALEXANDER WOODS

has been accepted towards fulfillment of the requirements for the

Ph.D.	degree in	ACCOUNTING AND INFORMATION SYSTEMS
Re-je-i	kish_c.	Man Jutot
	Major Pro	féssbr's Signature
	May	28, 30CF
	•	Date

MSU is an affirmative-action, equal-opportunity employer

PLACE IN RETURN BOX to remove this checkout from your record. TO AVOID FINES return on or before date due. MAY BE RECALLED with earlier due date if requested.

	DATE DUE	DATE DUE	DATE DUE
1			
		L	6/07 p:/CIRC/DateDue.indd-p.1

SUBJECTIVE ADJUSTMENTS TO OBJECTIVE PERFORMANCE MEASURES: AN EMPIRICAL EXAMINATION IN COMPLEX WORK SETTINGS

Bу

Alexander Woods

A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Accounting and Information Systems

ABSTRACT

SUBJECTIVE ADJUSTMENTS TO OBJECTIVE PERFORMANCE MEASURES: AN EMPIRICAL EXAMINATION IN COMPLEX WORK SETTINGS

By

Alexander Woods

This study examines the question: What factors determine whether supervisors subjectively adjust the objective performance measures of managers? I use survey and proprietary secondary data of 434 objective performance measures for 111 audit managers from 12 divisions of one of the world's largest internal audit organizations to study this question. The organization introduced a pay-for-performance incentive plan that allowed supervisors to adjust (up or down) the objective performance measures of audit managers. Empirical evidence documents both benefits and costs of subjectivity. Specifically, supervisors subjectively adjusted objective performance measures they perceived to be incomplete, unverifiable, and noisy, suggesting incentive contracting benefits. However, objective performance measures were also subjectively adjusted based on influence activities of managers and personal preferences of supervisors, suggesting incentive contracting costs.

Dedicated to the God of the universe and his son Jesus Christ for his unending love and amazing grace. To Him be the glory.

ACKNOWLEDGEMENTS

I owe much gratitude to my dissertation advisors Karen Sedatole (co-chair) and Ranjani Krishnan (co-chair). I greatly appreciate the caring, competent, fair, optimistic, and patient manner in which they provided guidance to me. Their selflessness made the completion of this dissertation possible.

I thank the other members of my dissertation committee, Sue Haka and Alex von Eye, for their insightful comments. I thank Edward Li, Fabienne Miller, Christian Mastilak, and workshop participants at Michigan State University for helpful comments on previous drafts as well.

I thank K. Ramesh for being a great leader of the PhD program and for always providing an encouraging word or two along the way.

Last but certainly not least, I am grateful for my wife Renee for her continual love, support and encouragement, my son Joshua for always making me laugh, and my daughter Hannah for helping to get my mind off all the work!

TABLE OF CONTENTS

LIST OF FIGURES	. vi
LIST OF TABLES	vii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW AND HYPOTHESES	6
2.1. Prior Subjectivity Research	6
2.2. Benefits of Subjective Performance Measure Adjustments	7
2.2.1. Performance Measure Sensitivity, Congruity, and Precision	8
2.2.2. Performance Measure Completeness	9
2.2.3. Performance Measure Verifiability	10
2.2.4. Performance Measure Manipulability	11
2.3. Costs of Subjective Performance Measure Adjustments	12
2.3.1. Influence Activities	13
2.3.2. Favoritism	14
	• •
CHAPTER 3: RESEARCH SETTING	16
3.1. Pay-for-Performance Incentive Plan	16
3.2. Data Collection and Survey Development	22
CHAPTER 4: VARIABLE MEASUREMENT AND EMPIRICAL SPECIFICATION.	24
4.1. Subjectivity	24
4.2. Benefits of Subjective Performance Measure Adjustments	25
4.3. Influence Activities and Favoritism	27
4.4. Control Variables	28
4.5. Empirical Specification	29
CHAPTER 5: RESULTS	30
5.1. Main Analysis	30
5.2. Sensitivity Tests	34
5.3. Supplementary Analysis and Discussion	34
5.4. Conclusion	36
CHAPTER 6: SUMMARY	37
APPENDICES	51
Appendix A: Review of the Literature on Subjectivity in Performance Evaluation and	d .
Compensation	<u>-</u> 52
Appendix B: Survey Instrument	60
	20
BIBLIOGRAPHY	73

LIST OF FIGURES

Figure 1: Organization Chart of Each Directorate	40
Figure 2: 2005 vs. 2006 Evaluations	41

LIST OF TABLES

Table 1: Analysis of Non-response Bias
Table 2: Summary of Procedures to Arrive at Final Performance Measure Sample 43
Table 3: Adjustments by Performance Measure, Division, and Directorate 44
Table 4: Common Factor Analysis with Varimax Rotation of the 12 PerformanceMeasure Characteristic Survey Items45
Table 5: Descriptive statistics of the 12 Performance Measure Characteristic Survey Items 46
Table 6: Descriptive statistics of the 4 Factors by Performance Measure 47
Table 7: Descriptive Statistics of Costs of Subjectivity and Control Variables 48
Table 8: Correlation Matrix of the Independent Variables
Table 9: Probit Regression of Subjective Performance Measure Adjustments 50
Table A1: Probit Model Including Downward Adjustments 65
Table A2: Ordered Probit Model Including Downward Adjustments
Table A3: Probit Model without Performance Measure Indicators 67
Table A4: Probit Model with Alternative Division Performance Variable 68
Table A5: Probit Model with Alternative Level Variable 69
Table A6: Probit Model with Alternative Adjustment Variable
Table A7: Probit Model using Huber-White Cluster-Corrected Standard Errors
Table A8: Probit Model with Number of Managers as an Additional Control Variable 72

CHAPTER 1: INTRODUCTION

In this dissertation, I investigate the question: What factors determine whether supervisors subjectively adjust the objective¹ performance measures of managers? This is an important question because the use of subjectivity in performance evaluations is widespread and increasingly viewed as appropriate in a knowledge-based economy (MacLeod and Parent 1999; Haag et al. 2006). Specifically, "knowledge workers" perform jobs that require them to allocate their effort across complex and multidimensional work tasks (Holmstrom and Milgrom 1991). Managerial work is inherently complex because it requires managers to (a) identify and focus on only the most critical data that enables them to define and understand relevant issues, (b) decompose and/or integrate complex data, (c) determine connections between disparate pieces of data to detect emerging patterns and trends, and (d) identify and evaluate alternative, creative, insightful, and logical solutions when a situation is ill-defined and there is a high degree of uncertainty about the outcome (Ilgen and Hollenbeck 1991). As a result of the complexity, it is difficult to specify all relevant dimensions of manager effort and sometimes even desirable outcomes of that effort. Accordingly, objective performance measures used to evaluate managers can fail to indicate a manager's contribution toward organizational objectives, induce distortions in managerial effort, and provide managers with ineffective and/or costly incentives (Hopwood 1972; Milgrom and Roberts 1992; Baker, Gibbons, and Murphy 1994; Prendergast 1999; Baker 2000). In other words, although the extensive use of objective measures suggests their benefits outweigh their costs, these measures are nevertheless often deficient (i.e., imperfect).

¹ I use objective to mean quantitative performance information.

However, supervisors observe informative but noncontractible information about managerial performance not captured by the available objective performance measures. Thus, organizations commonly allow supervisors to subjectively adjust objective performance measures to compensate for deficiencies in the measures in order to obtain a more accurate representation of a manager's contribution toward organizational objectives, mitigate distortions in managerial effort, and improve incentive contracting (Merchant 1989; Baker et al. 1994; Baiman and Rajan 1995; Murphy and Cleveland 1995; Fisher et al. 2005). Although organizations can reap benefits when supervisors subjectively adjust objective performance measures for their deficiencies, empirical evidence shows that subjective performance evaluations and subjective performance measure weighting decisions often exhibit biases (Murphy and Cleveland 1991; Arvey and Murphy 1998; Krishnan et al. 2005; Orey 2006). These biases occur in part because of the multi-layered relationships that actually exist in most hierarchical organizations where middle-level supervisors evaluate lower-level managers and receive only partial and/or indirect benefits from manager productivity (i.e., supervisors are not often full residual claimants on manager output). Thus, supervisors' incentives with respect to their subjective performance measure adjustments are often imperfectly aligned with organizational objectives (Prendergast and Topel 1993; Prendergast 2002; Bol 2005). Accordingly, it is not clear whether supervisors subjectively adjust objective performance measures to compensate for deficiencies in the measures as suggested by economic theory or because of social influences that result from the social context in which subjective adjustments occur that standard principal-agent models do not consider.

Data limitations usually prevent the study of subjective performance measures adjustments. Specifically, researchers cannot normally observe adjustments made to arrive at measurement outcomes because only final measurement outcomes are recorded. However, I collected secondary data from a proprietary performance evaluation and compensation plan for audit managers which recognizes that manager performance on each objective measure in the plan may not appropriately represent the manager's performance. Accordingly, the plan includes a "line item" that allows supervisors to subjectively adjust manager performance on each objective performance measure if/when supervisors perceive that performance as indicated on the measure is not representative of managers' contributions to the organization's mission and goals.

I combine this proprietary secondary data, collected from a pay-for-performance incentive plan introduced by one of the world's largest internal audit organizations, with survey data that measures perceptions about the properties of the objective performance measures and the way in which supervisors adjusted the measures. The empirical findings show both benefits and costs of subjective performance measure adjustments. On one hand, supervisors subjectively adjusted performance measures to compensate for measures they perceived to be incomplete, unverifiable, and noisy, suggesting incentive contracting benefits. In addition, consistent with prior theoretical research (Holmstrom and Milgrom 1991), supervisors subjectively adjusted performance measures that were intended to capture more of the task multidimensionality inherent in complex jobs. Further, the introduction of pay-for-performance significantly mitigated the most common performance evaluation biases, leniency and compression, as indicated by the reduction in the skewness and kurtosis of the evaluation distribution. On the other hand,

the empirical evidence also documents that supervisors subjectively adjusted performance measures for influence activities of managers and for supervisory favoritism, suggesting incentive contracting costs. In addition, subjective adjustments were not related to performance measure sensitivity, congruity, or manipulability, suggesting that not all of the potential benefits of subjectivity were realized.

This dissertation contributes to the accounting literature in at least five ways. First, this study provides distinctive insight into performance measurement in subjective evaluation and compensation systems because subjective adjustments are made to each of the performance measures within managers' bonus contracts. As such, it extends prior research findings about the effect of formula bonus contract deficiencies on the use of subjective bonus awards (Gibbs et al. 2004). Second, by analyzing the use of subjectivity for managers who perform complex and multidimensional tasks (the most ideal setting to examine the use of subjectivity because objective measures of value-enhancing efforts for such managers are difficult to obtain) (Holmstrom and Milgrom 1991; Prendergast 1999), this dissertation contributes to our understanding of how task complexity may affect supervisory perceptions of the economic properties of performance measures. For example, when managers perform complex vis-à-vis simple tasks (Gibbs et al. 2004; Krishnan et al. 2005), it is conceivable that the noisiness of a measure is not only more salient to supervisors, but is also excessive (i.e., their perception of the level of noisiness is beyond the true level of noisiness). Third, this study enriches and expands theoretical development of subjectivity research by identifying that different forms of subjectivity can have opposing predictions, and by highlighting and providing empirical evidence of the importance of social costs that accompany the use of subjectivity.

Fourth, this study provides important insights about compensation to internal auditing practitioners and researchers because the nature of internal auditing work is becoming more complex and internal auditing functions are increasing in size as a result of the Sarbanes-Oxley Act of 2002 (McDonald 2003), and because the literature has acknowledged the need for such research (Prawitt 2003). Finally, by considering social, cognitive, and economic effects of the use of subjectivity, my study answers calls for, and contributes to, accounting research aimed at integrating behavioral and economic factors (e.g., Waller 1995; Kachelmeier 1996; Evans, Hannan, Krishnan, and Moser 2001; Merchant, Van der Stede, and Zheng 2003).

The remainder of the dissertation is organized as follows. Chapter II formalizes the research hypotheses. Chapter III describes the research setting. Chapter IV describes the measurement of the variables and the empirical specification used to test the hypotheses. Chapter V presents the results, and Chapter VI provides a summary discussion of the results including their limitations and implications.

CHAPTER 2: LITERATURE REVIEW AND HYPOTHESES

2.1. Prior Subjectivity Research

Prior literature has studied three forms of subjectivity (see Appendix A): individual performance evaluations, discretion in weighting objective performance measures, and paying bonuses based on factors other than the objective measures specified in managers' bonus contracts (e.g., Govindarajan 1984; Baker et al. 1994; Ittner et al. 2003; Gibbs et al. 2004). First, subjective (individual) performance evaluations are used to reduce noise in performance evaluations and reduce distortions in CEO effort (Bushman et al. 1996). Second, subjective performance measure weights are affected by whether performance measures are financial vs. nonfinancial, internal vs. external, objective vs. subjective, and outcome vs. driver (Ittner et al. 2003). Finally, subjective bonus awards are positively related to departmental interdependencies, the achievability of formula-based bonuses, financial losses, and long-term investments in intangibles, indicating that subjective bonus awards are used to compensate for deficiencies in the incentive contracts of managers with less complex jobs (Gibbs et al. 2004). These and other subjectivity studies provide important insight into the use of subjectivity in compensation contracts, but do not examine whether subjectivity is used for the primary benefit identified by economic theory, namely, to compensate for deficiencies in objective measures (Baker et al. 1994; Prendergast 1999; Murphy and Oyer 2003). Studying the different forms of subjectivity is important because the different forms can

have opposing predictions, thereby enriching theoretical development of subjectivity research.²

In this chapter, I consider whether supervisors are likely to subjectively adjust objective performance measures when they perceive the measures are deficient with respect to performance measure properties specified by economic theories of performance measurement. Then, I draw on economic, social, and cognitive theories to consider whether subjective performance measure adjustments could be impacted by influence activities of audit managers and personal preferences of supervisors.

2.2. Benefits of Subjective Performance Measure Adjustments

Subjective performance measure adjustments can improve incentive contracting because they allow supervisors to exploit the private information they observe about manager performance that is not represented by the objective performance measures specified in the manager's incentive contract (Hayes and Schaefer 2000; Gibbs et al. 2004; Fisher et al. 2005; Van der Stede, Chow, and Lin 2006). Objective performance measures are most likely to be deficient when managers perform multidimensional and complex tasks because these tasks are difficult to measure (Hopwood 1972; Holmstrom and Milgrom 1991; Prendergast 1999). Prior research indicates that performance measures are deficient when they are (a) not sensitive to manager actions, (b) not congruent with organizational objectives, (c) noisy, (d) incomplete (e) not verifiable, and (f) manipulable (Banker and Datar 1989; Holmstrom and Milgrom 1991; Feltham and

² For example, prior research documents that subjective performance measure weights are increasing in performance measure verifiability (Ittner et al. 2003). If subjective performance measure weights are a substitute for subjective performance measure adjustments, then subjective adjustments should also be increasing in verifiability. However, this dissertation predicts and finds that more verifiable measures are less likely to be subjectively adjusted.

Xie 1994; Baker et al. 1994). The more that objective performance measures are deficient, the more that subjective adjustments to performance measures can be used to properly represent manager performance, mitigate distortions in managerial effort, and improve incentive alignment.

2.2.1. Performance Measure Sensitivity, Congruity, and Precision

Agency theory research identifies three important properties of objective performance measures: sensitivity, congruity, and precision (Holmstrom 1979; Banker and Datar 1989; Feltham and Xie 1994; Datar, Kulp, and Lambert 2001). A measure is more *sensitive* when the agent's action has a larger expected effect on it. A measure is more *congruent* when an agent's action that increases a measure also increases the principal's gross payoff. A measure exhibits more precision (less noise) when uncontrollable causes explain less of the variation in the measure. The extent to which a performance measure exhibits these three properties determines the extent to which that performance measure is deficient and the extent to which an adjustment to it would improve incentives. Specifically, objective measures that are low in sensitivity, congruity, and precision are less informative about underlying managerial performance. By subjectively adjusting uninformative measures to compensate for the deficiencies in the measures, supervisors can make the measures more informative about underlying managerial performance and improve incentive contracting. In contrast, objective measures that are high in sensitivity, congruity, and precision need not be subjectively adjusted because they are already informative about the underlying managerial performance for which they were intended to capture. Thus, I present the following hypotheses:

HYPOTHESIS 1. The likelihood of supervisors' subjective adjustments to objective performance measures will be negatively (positively) related to supervisors' perceptions of performance measure sensitivity (insensitivity). HYPOTHESIS 2. The likelihood of supervisors' subjective adjustments to objective performance measures will be negatively (positively) related to supervisors' perceptions of performance measure congruity (incongruity). HYPOTHESIS 3. The likelihood of supervisors' subjective adjustments to objective performance measures will be negatively (positively) related to supervisors' perceptions of performance measure congruity (incongruity).

2.2.2. Performance Measure Completeness

Complex jobs involve managers allocating their effort across numerous task dimensions. Some of those dimensions are outcomes and some are detailed activities that give rise to the final outcome (Hopwood 1972). A single objective performance measure may not adequately capture all task dimensions. As a result, multiple performance measures are often used to evaluate managers working on multidimensional and complex tasks (Kaplan and Norton 1996, 2001; Lipe and Salterio 2000). Ideally, each measure in the optimal incentive contract is incrementally informative about the manager's efforts across required task dimensions (Holmstrom 1979; Antle and Demski 1988). However, even the use of multiple objective measures may not capture all of the important dimensions of the manager's job. As a result, compensation contracts are often incomplete. This can give rise to distortions in managerial effort because managers have

the incentive to exert effort only on the dimensions of performance that are captured by the objective performance measures in their incentive contract (Holmstrom and Milgrom 1991). Subjective performance measure adjustments can be used to provide proper managerial incentives when the objective measures in the manager's contract fail to account for all the important dimensions of a manager's task performance. However, if the objective performance measures capture the important dimensions of a manager's performance, then subjectively adjusting such objective performance measures provides little benefit in the form of improved incentive contracting. Thus, I predict that:

HYPOTHESIS 4. The likelihood of supervisors' subjective adjustments to objective performance measures will be negatively (positively) related to supervisors' perceptions of performance measure completeness (incompleteness).

2.2.3. Performance Measure Verifiability

Another property of a performance measure that is important in subjective evaluation systems is verifiability, which is the extent to which a measure can be easily quantified and/or specified (Rajan and Reichelstein 2006; Moers 2006). Because manager evaluations often have consequences for supervisors, supervisors are likely to consider how manager evaluations affect their own performance evaluations (Ilgen and Favero 1985). Supervisors who adjust "harder" measures that can be easily quantified and verified can invite accusations of misconduct such as favoritism, an often-cited cost of subjectivity (Prendergast and Topel 1993, 1996). On the other hand, "softer" measures that are less amenable to quantification and/or verification can enable supervisors to

justify the adjustment of measures with less fear of such accusations. Accordingly, I expect adjustments to objective performance measures to be less likely with increases in performance measure verifiability. This idea is formalized in the following hypothesis:

HYPOTHESIS 5. The likelihood of supervisors' subjective adjustments to objective performance measures will be negatively (positively) related to supervisors' perceptions of performance measure verifiability (un-verifiability).

2.2.4. Performance Measure Manipulability

Some objective performance measures are more susceptible to manipulation by managers than other measures. Specifically, managers may be able to manipulate objective performance measures so that the measures indicate performance has improved without actually producing a *real* performance improvement (Baker 1990; Baker et al. 1994). For example, mechanics at Sears increased profits by misleading customers into authorizing unnecessary repairs (since the mechanics were paid commissions based on the profits from repairs authorized by customers), resulting in Sears closing its auto repair business in California (Baker et al. 1994). Allowing supervisors to subjectively adjust and "back out" the effects of manipulated performance measures lessens managers' incentives to manipulate the measures and increases managers' incentives to engage in organizationally productive activities. Further, having previously been managers themselves, supervisors know how managers can manipulate performance on a measure without affecting real performance. This line of reasoning leads to the following hypothesis:

HYPOTHESIS 6. The likelihood of supervisors' subjective adjustments to performance measures will be positively related to supervisors' perceptions of performance measure manipulability.

2.3. Costs of Subjective Performance Measure Adjustments

While the use of subjectivity can provide supervisors with the latitude to adjust objective performance measures for their deficiencies, resulting in incentive contracting benefits, the use of subjectivity also entails costs that can impair incentive contracting (Moers 2005; Bailey, Hecht and Towry 2006). Specifically, prior research shows that when making subjective performance measure weighting decisions, individuals often do not adjust performance measure weights in the direction predicted by agency theory in response to changes in the precision and error covariance of performance measures (Krishnan et al. 2005). In addition, individuals weight measures that are common across divisions more in evaluating division performance than they weight measures that are unique across divisions because of cognitive limitations (Lipe and Salterio 2000).

While these and other studies provide important insights into cognitive biases associated with subjective performance measure weighting decisions, they do not examine the social context of subjectivity in performance evaluations (Longenecker, Sims, and Gioia 1987). In particular, the use of subjectivity in multi-layered hierarchical organizations grants one agent (supervisors) power over the welfare of other agents (managers). As a result, managers have an incentive to spend time trying to influence their supervisor's subjective assessment of their performance instead of trying to accomplish organizationally productive activities. Similarly, supervisors have the

freedom to act on personal preferences toward managers and reward some managers over others (i.e., exhibit favoritism). In this chapter, I consider whether manager influence activities and supervisor favoritism will be associated with supervisors' subjective performance measure adjustments.

2.3.1. Influence Activities

One cost of subjectivity is the development of upward influence activities and tactics, or ingratiatory behavior (Milgrom and Roberts 1988). Specifically, because managers are importantly affected by their supervisor's discretionary use (and non-use) of subjective adjustments, managers have an incentive to spend time trying to influence their supervisor's subjective assessment of their performance (Ferris and Judge 1991). This is costly because engaging in activities designed to favorably influence supervisory assessments diverts managers' attention and effort from more organizationally productive activities (Anderson et al. 2007).

How will supervisors respond to a manager's influence tactics? Supervisors might respond by downwardly adjusting the manager's performance to refocus manager effort on organizationally productive tasks. Or, supervisors might dismiss the manager's behavior entirely. However, there are two countervailing reasons to expect manager influence activities to successfully sway supervisors' subjective adjustments. First, prior research has shown that, in many settings, managers successfully negotiate their performance standard downward (i.e., create "slack"), suggesting they spend significant (and productive) time trying to influence their supervisors instead of increasing their effort on job performance (Argyris 1964; Fisher et al. 2000, 2002; Anderson et al. 2007). Second, managers' influence tactics favorably sway supervisors' perceptions of

managers' promotability and skills (Wayne et al. 1997). Thus, prior research suggests that managers will engage in, and be successful at, influence activities. This discussion leads to the following hypothesis:

HYPOTHESIS 7. The likelihood of supervisors' subjective adjustments to objective performance measures will be positively related to perceptions of higher manager influence activities.

2.3.2. Favoritism

While subjectivity is expected to introduce distortions in effort that arise from influence activities of *managers*, it is also expected to introduce distortions that arise from biases of the *supervisor*. Specifically, an often-cited cost of the use of subjectivity in performance evaluation systems is favoritism, where supervisors reward some managers over others based on supervisors' personal preferences toward or relationships with managers (and not based on managers' performance) (Prendergast and Topel 1996;). It is important to note that while there is a lot of evidence of complaints of perceived favoritism (e.g., Ittner et al. 2003), there is little evidence to support its actual existence.

Favoritism occurs in part because of the hierarchical relationships that exist in most organizations where supervisors evaluate managers but are not full residual claimants on manager output (i.e., the benefits that supervisors receive from manager productivity are often *partial* and/or indirect) (Baker et al. 1994; Fisher et al. 2005; Prendergast 2002). In other words, supervisors' incentives with respect to their evaluation of managers are often imperfectly aligned with organizational objectives. Favoritism can

have significant economic consequences to an organization because it results in ineffective and inefficient incentive contracting, arbitrary reward assignments, and less productive job assignments (Prendergast and Topel 1996). In addition, because it is both a cause and a consequence of organizational politics and power struggles, favoritism can have further indirect spillover effects on motivation and productivity through its detrimental impact on organizational conflict and stress (Kwon 2005). I posit the following formal hypothesis related to favoritism:

HYPOTHESIS 8. The likelihood of supervisors' subjective adjustments to objective performance measures will be positively related to perceptions of higher supervisor favoritism.

CHAPTER 3: RESEARCH SETTING

3.1. Pay-for-Performance Incentive Plan

The research setting for this dissertation is 12 divisions of a large internal audit organization. Prior to 2006, this organization provided its audit managers (AMs) with annual salary increases based on seniority. A nominal bonus was also provided based on an overall subjective evaluation of performance on a scale from 9 to 81. As part of a comprehensive "performance management plan" aimed at recruiting, retaining, and rewarding talented human capital, the organization introduced a pay-for-performance incentive plan in 2006 and made managers' annual salary increase and bonus dependent on managers' performance as indicated on four different measures of performance. The four measures were chosen because they capture different dimensions of managerial performance that the organization believed would represent managers' contributions to organizational objectives.

The Planning performance measure instills ownership over future audit work. Specifically, this measure requires managers to annually identify audit subjects and contribute to the development of a risk-based plan that is focused on improving the organization's processes. Success on this objective requires managers to integrate knowledge and technical proficiency so that customer needs are best served in the future.

The Program Management measure captures the bulk of audit managing work that occurs during the performance measurement period. Program management requires audit managers to perform preliminary research, program design, application, summarization and analysis work meeting specified requirements of the organization, the organization's

strategic plan, and generally accepted government auditing standards. Managers must identify resource requirements prior to the application phase of the audit. They must provide technical guidance and assistance to field auditors during the application phase of the audit to facilitate timely audit completion. Managers must also prepare work papers that include sufficient evidence to support audit objectives and conclusions without material omissions. Managers must proficiently use computerized techniques to enhance the audit process and analyze collected audit data. They also inform supervisors immediately of all critical changes to the audit plan, such as milestone delays, audit application problems, and customer requests.

The Reporting measure captures the end product of audit manager work. Specifically, the final result of an audit is a unique audit report (or other product such as outlines, drafts, or briefings) which presents audit results, identifies causes, and recommends corrective actions. The Reporting measure requires audit managers to effectively communicate by writing audit reports that require minimal revisions. It also requires managers to present and defend balanced and objective audit findings and recommendations to organization management officials, keeping customers informed and appropriately considering customer views and disagreements.

The Professional Development measure ensures audit manager competency in their respective area of expertise. It requires audit managers to identify training opportunities for continuing education and professional development. The organization believes the acquisition and use of job-specific knowledge drives organizational performance. As a result, managers must complete functional training and meet prespecified education requirements in accordance with organization standards.

Each of the performance measures were objectively determined and quantified as follows. First, performance standards for each of the four objective performance measures were established prior to the measurement period. Then, at the end of the measurement period, managers' performance was recorded on a scale from one to five in accordance with the pre-established standards. For example, for the Planning performance measure, managers were expected to identify and document two audit subjects for inclusion in the organization-wide triennial audit plan. If managers identified and documented two audit subjects that were included in the organization-wide audit plan, they would receive a "3" on the Planning measure. But if managers contributed four audit subjects to the audit plan, they would receive a "5" since their performance on this measure was well beyond the established standard. Or if managers contributed less than two audit subjects, a score of less than "3" would be recorded for their performance on this measure. Similarly, for the Program Management performance measure, an audit plan that detailed assigned tasks, milestones, and budgeted hours for every phase of the audit was negotiated and agreed upon. If managers accomplished assigned tasks within agreed-to milestones and budgeted hours 90 percent of the time, they would receive a "3" on the Program Management measure. However, if they achieved results above (below) this standard, then a score higher (lower) than "3" would be recorded for their performance on this measure.

The organization recognized the potential for deficiencies in each of these objective performance measures however, and granted supervisors discretion to adjust

performance on each measure one unit up or down to compensate for the deficiencies.³ This resulted in an adjusted measure. Performance on the four adjusted measures was averaged to determine an overall "average score" between one and five.⁴ The overall average score automatically determined the total amount of incentive pay managers received. For example, one manager received a 3, 3, 3, and 3 on the objective performance measures. That manager's supervisor then adjusted the measures 0, +1, +1, and +1, resulting in adjusted scores of 3, 4, 4, and 4, respectively. The manager's overall average score was 3.75 [(3 + 4 + 4 + 4)/4]. The manager was allocated \$3,290 of incentive pay, \$1,304 more than the mean incentive pay of \$1,986.

Adjustments are important to managers not only because they affect managers' incentive pay but also because they affect managers' future promotion potential. The authority given to supervisors to adjust the measures, as well as all other aspects of the new pay plan (e.g., the measures, the pre-determined criteria, etc.), was transparent to both managers and supervisors. It is also important to note that supervisors' pay plans included five equally-weighted objective performance measures, one of which related to their ability to develop the managers under their supervision.

Although supervisors can subjectively adjust the objective performance measures in managers' performance plan, the organization also uses a "leveling" mechanism as part of the performance evaluation and compensation system. Specifically, after supervisors have signed off on their evaluations of managers, each of the 12 division

³ For example, organizational documents discuss how adjustments help supervisors to beneficially "align individual work with (the organization's) mission and priorities" and can be used to "account for factors that are necessary for effective, efficient work accomplishment".

⁴ The "template" of the pay plan allowed unequal weighting of the four performance measures. However, the organization's senior leaders mandated equal weighting. Therefore, subjective weighting of the measures was not under supervisors' discretion.

managers (supervisors of the supervisors) reviews and can override supervisors' evaluations of managers within the division. Division managers perform this performance calibration procedure to ensure all audit managers in the division are fairly and similarly evaluated across different supervisors. Each division belongs to one of three directorates; each directorate is a separate pay pool and is geographical separated from the other directorates. See Figure 1 for an organization chart of each directorate.

Although the same set of performance measures and standards are being used to evaluate similar managers doing "internal auditing," each of the audit divisions perform audits that differ on numerous dimensions. Thus, the jobs differ with respect to their complexity, and with respect to other important aspects such as the required knowledge and skills of managers. For example, the Financial Management division performs mandated financial statement audits (the organization does not discretionarily choose to perform these audits). Managers in this division require knowledge related to financial accounting and successful audits are based on timely meeting reporting requirements. Conversely, the Maintenance division discretionarily chooses the types of maintenance audits it will perform. The knowledge required in this division relates to the organization's maintenance activities and successful audits often indicate a dollar amount of organizational savings the audit has identified. As a result, the degree to which each of the four objective performance measures are sensitive, congruent, noisy, complete, verifiable, and manipulable varies with respect to division.

This setting is ideal to examine whether and why supervisors subjectively adjust objective performance measures of audit managers to compensate for deficiencies in the measures. In particular, the organization's pay-for-performance plan explicitly recognizes

the potential for deficiencies in the objective performance measures in the plan. In addition, the proprietary secondary data enables me to directly measure subjective performance measure adjustments based on deficiencies in the objective measures; data for such direct measurements were not available from prior studies. Further, the organization did not allow supervisors to use other forms of subjectivity in the evaluation and compensation process. Specifically, the organization did not allow supervisors to (1) subjectively weight the objective measures, (2) subjectively decide the amount of incentive pay to distribute to managers, or (3) use subjective performance evaluations. Thus, I am implicitly controlling for the impact of other forms of subjectivity that have been studied in prior research (Ittner et al. 2003). Finally, the same set of performance measures and standards were used to evaluate similar managers doing "internal auditing." This homogeneity means that I am implicitly controlling for many variables that would otherwise be difficult to control for.

However, I have no reason to believe that my results are specific to this one organization. The hypotheses are not specific to federal organizations, internal audit organizations, or even more generally to accounting and/or auditing organizations; rather, the hypotheses generalize to many organizational settings in which tasks are complex and subjective components are used to supplement objective measures of managerial performance.

3.2. Data Collection and Survey Development

I collected data from several sources to test the hypotheses. First, I collected various division performance metrics directly from the internal audit organization.⁵ Second, I collected the proprietary evaluation and compensation data (i.e., performance on the objective measures, adjustments to the measures, final score on the measures, and the resulting incentive payout for all 130 managers in the organization's 12 divisions for the 2006 evaluation year) via a Freedom of Information Act (FOIA) request from the federal organization with authority over the internal audit organization's implementation of this pay-for-performance plan.⁶ Third, I developed three surveys directed at each management layer in the organization hierarchy: (1) managers, (2) supervisors, and (3) division managers. The surveys asked respondents about their perceptions regarding (a) the properties of the objective performance measures used in the performance incentive plans of managers, (b) the way supervisors used adjustments, (c) manager job complexity and influence activity, (d) the information asymmetry between supervisors and division managers about the performance of managers, and (e) satisfaction with the new pay plan. The supervisor and division manager surveys also asked questions about the extent to which respondents believe division managers adjusted supervisors' evaluations of managers, and the manager survey asked questions about fairness in the evaluation and compensation system. (See Appendix B for the survey instrument questions related to

⁵ The organization also provided evaluation and compensation data that includes overall ratings for years 2004 through 2006. This information identifies the manager whose performance was evaluated, the supervisor who evaluated the performance of each manager, and the division to which each supervisor/manager pair belong.

⁶ The organization that I obtained the proprietary data from is overseeing the pay-for-performance plan described in this paper for more than 100,000 employees across more than 30 organizations. Thus, the pay plan is not unique and the results of this study provide informative analyses specifically relevant to other organizations under the same plan and, more generally, to other organizations introducing subjective elements into their performance evaluation and compensation system.

performance measure properties, job complexity and influence activity, and the use of adjustments).

The surveys were conducted using Internet survey software. Emails were sent to the administrative assistants of each of the three directorates, who then forwarded separate emails to division managers, supervisors, and managers in the directorate. The email described the purpose of the survey, included the survey links, and informed respondents that the survey links would be available between March 5 and 23, 2007. I obtained survey responses from 8 of the 12 division managers (67 percent), 20 of the 36 supervisors (56 percent), and 39 of the 130 managers (30 percent). The distribution of responses across divisions appears that all divisions were adequately represented. Specifically, the mean number of responses per division was 5.6 and ranged between 4 and 9 (standard deviation of 1.4).

I compared the respondents and non-respondents on variables available for both to assess the possibility of non-response bias. The mean number of shares distributed (2.41), mean incentive payout (\$1,986), and the percentage of manager respondents with at least one adjusted performance measure (63 percent) was not statistically different than the means for manager non-respondents (p > .05, two-tailed). This leads me to believe that respondents do not differ in important ways from non-respondents. See Table 1 for the non-response bias analysis.

CHAPTER 4: VARIABLE MEASUREMENT AND EMPIRICAL

SPECIFICATION

In this chapter, I describe the measurement of the variables and the empirical specification used to test the hypotheses in Chapter 2.

4.1. Subjectivity

My hypotheses examine factors that affect whether supervisors subjectively adjust objective performance measures. The unit of analysis is the subjective adjustment to an individual performance measure for a given audit manager. Thus, the proprietary evaluation data was comprised of 4 performance measures for each of 111 managers, for an available sample of 444 performance measures. One hundred and thirty-eight (31 percent) of these performance measures were adjusted, 93 percent (128/138) of which were adjusted upward (in a manner that increased performance on the performance measure).⁷ Because this small number of downward adjustments precludes statistical analyses that differentiate between upward and downward adjustments, I deleted the 10 downwardly adjusted measures and measured subjectivity as the *presence* of an upward adjustment.⁸ Table 2 describes how the final sample of 434 measures was derived. Table 3 provides descriptive statistics of the adjustments by direction, directorate, division, and performance measure.

⁷ This finding may be explained by supervisors' objective function. Specifically, one of the objective performance measures in their pay plan evaluates their ability to develop the managers under their supervision. Thus, downward adjustments may imply an admission by supervisors that they were not effective on that performance dimension.

⁸ Results are qualitatively similar when I do not delete the 10 downwardly adjusted measures.

4.2. Benefits of Subjective Performance Measure Adjustments

Using twelve questions adapted from Moers (2006), Gibbs et al. (2004), and Anderson et al. (2007), I surveyed supervisors and division managers to measure their perceptions of the extent to which each of the four performance measures were deficient.⁹ A seven-point, fully-anchored scale was used to indicate the extent to which each performance measure reflected the following properties: sensitivity, congruity, noise, completeness, verifiability, and manipulability (see Appendix B for the questions used for each property). Thus, each respondent was asked 48 questions related to performance measure properties (12 questions for each of 4 performance measures).

I constructed the performance measure property variables in two steps. In the first step, I performed common factor analysis with varimax rotation on the set of 12 questions to identify the underlying properties of performance measures *as perceived by the supervisors and division managers* and to ensure discriminant validity across the constructs. Because an analysis of the distributions indicated that supervisor and division manager perceptions of performance measure properties did not differ, and to mitigate the effects of missing survey responses across divisions, I pooled the responses from supervisors and division managers for purposes of the factor analysis.¹⁰ In addition, because the survey questions provide equivalent (i.e., "similarly interpretable") measures of the underlying properties of the four performance measures, I also pooled the data

⁹ I actually asked respondents 14 questions. The two additional questions were about whether the performance measures captured the quality and quantity of manager performance. However, I did not use these questions in subsequent analyses because various factor analyses revealed that they did not correlate highly with the other questions to which they were a factor. Therefore, I do not discuss these questions.

¹⁰ I also surveyed audit managers for comparison. However, the perceptions that matter are those of supervisors and division managers since they are the ones responsible for adjusting the performance measures. Interestingly, audit managers perceived three of the four performance measures were significantly more deficient than supervisors and division managers with respect to all but one performance measure property (with the specific property being different for each measure).

across the four performance measures, resulting in a final factor analysis sample of 112 cases (28 supervisor and division manager responses x 4 performance measures) for the 12 survey questions.

In the second step of the process, I aggregated the factor scores by division to generate a mean factor score (for each performance measure property construct for each performance measure) to be used in the main analysis. In one instance, I had only one factor score (the division manager response). The resulting factor analysis identified four factors with eigenvalues greater than one: (1) sensitivity/congruity, (2) completeness/verifiability, (3) noise, and (4) manipulability.¹¹ The four factors explained 77 percent of the total variance, their factor loadings ranged from 10.571 to 10.921, and Cronbach's alpha was greater than 0.80 for all factors. See Table 4 for the twelve statements, the four factors, and their factor loadings and eigenvalues; see Table 5 for descriptive statistics of the individual questions; see Table 6 for descriptive statistics of the factor scores used in the main analyses.

It is noteworthy that sensitivity and congruity loaded on the same factor because these are theoretically distinct concepts. Treating these concepts as separate in the analyses that follow however would violate construct validity because various tests of discriminant validity all consistently revealed one underlying latent construct. There are three possible explanations: (1) supervisors and managers in this organization cannot distinguish between sensitivity and congruity, (2) sensitivity and congruity are highly correlated in this organization, and/or (3) my survey questions failed to distinguish sensitivity from congruity. The theoretically distinct performance measure properties of

¹¹ I conducted the factor analyses using different rotation methods. Results were qualitatively similar with all methods. I used common factor analysis with varimax rotation for my final factor analysis because it yielded the most interpretable factor structure.
verifiability and completeness also capture one underlying construct in this organization. This is not too surprising however, because other performance measures simultaneously exhibit both of these properties as well. For example, profit is a measure well-recognized as capturing numerous dimensions of task performance while also being verifiable.

4.3. Influence Activities and Favoritism

Incentive contracting can be costly if supervisors subjectively adjust objective performance measures for manager influence activities and supervisor favoritism. To measure these costs of subjectivity, the survey contained four questions (two questions each) adapted from Tziner et al. 1996 that had seven-point, fully-anchored scales. To measure influence activity, I asked managers, supervisors, and division managers to indicate the extent to which managers in the division try to favorably influence supervisors' (1) assessment of manager performance and (2) impression of managers. A common factor analysis revealed these two items were one factor (Cronbach's alpha was .87). I summed and averaged responses to each question by division, and then combined and averaged the two questions to create a composite InfluenceActivity variable. To measure favoritism, I asked managers and supervisors to indicate the extent to which supervisors made adjustments to managers' most recent performance evaluation to (1) increase rewards for some managers relative to other managers under the supervisor's supervision and (2) to improve the quality of the supervisor's personal relationship with the manager. A common factor analysis revealed these two items were one factor (Cronbach's alpha was .56). I summed and averaged responses to each question by

27

division, and then combined and averaged the two questions to create a composite *Favoritism* variable. See Table 7 for descriptive statistics of these variables.

4.4. Control Variables

In the final sample of 434 manager-performance measure observations, I include indicator variables for each performance measure to control for performance measure characteristics other than those measured. Similarly, to control for division differences (e.g., task complexity, audit quality) that may be correlated with perceptions of performance measure properties, I include indicator variables for each division (Div_i), and a measure of division performance (DivisionPerformance) that captures the average quality of audit reports produced by each division. Each final audit report is scored from 1 to 20 on a number of pre-determined dimensions important to the organization's primary customer. Division performance is important because it should correlate with adjustments of performance measures as well as other division practices. In addition, this measure of performance is a comprehensive measure that should capture performance differences across the divisions as indicated on the objective performance measures of managers. Finally, I use the pre-adjusted level of each objective measure as a control variable (Level) because there could be a non-empirical negative association built into the objective measure scale. For example, measures scored as a 5 can only be adjusted downward. In addition, an adjustment from 2 to 3 could imply a qualitatively different meaning of performance than an adjustment from 4 to 5.

28

4.5. Empirical Specification

To test the hypotheses, I estimate the following probit model using the variables identified above:

$$\begin{split} A_i &= \beta_0 + \beta_1 Sensitivity/Congruity_i + \beta_2 Noise_i + \beta_3 Completeness/Verifiability_i \\ &+ \beta_4 Manipulability_i + \beta_5 InfluenceActivity_i + \beta_6 Favoritism_i + \\ &\beta_7 Planning_i + \beta_8 Reporting_i + \beta_9 ProfessionalDevelopment_i + \\ &\beta_{10-17} Div_i + \beta_{18} DivisionPerformance_i + \beta_{19} Level_i + e_i \end{split}$$

where a subjective adjustment to the performance measure (A) is my dependent variable and the different performance measure properties, costs of subjectivity, and control variables are my independent variables. Subjective adjustments take on a value of 0 or 1, indicating the performance measure was not adjusted or it was adjusted, respectively. The Program Management performance measure indicator and five division indicators are the omitted (i.e., reference) variables.

CHAPTER 5: RESULTS

5.1. Main Analysis

The probit analysis in Table 9, Model 2 correctly classified 72.12 percent of the subjective performance measure adjustments (pseudo r-square of 10.9 percent).¹² My hypotheses make predictions about whether supervisors subjectively adjusted deficient objective performance measures. H1 and H2 predicted that supervisors' subjective adjustments to performance measures would be negatively related to supervisors' perceptions of performance measure sensitivity and congruity, respectively. Because the factor analysis reported in Chapter III revealed one underlying latent construct for sensitivity and congruity, the coefficient on the Sensitivity/Congruity variable in Table 9 is a joint test of both H1 and H2. The coefficient is not significant (-0.099, p > .10). Thus H1 and H2 are not supported. This is an interesting finding because the context of this study was ideal for examining whether supervisors subjectively adjusted performance measures for deficiencies in these specific characteristics. I offer three explanations for this non-significant result. First, it may be that productive agent actions and effort that move performance measures (sensitivity) toward fulfilling organizational goals (congruity) are indistinguishable to supervisors or are just not as cognitively salient or perceptible as the effect of uncontrollables impacting a performance measure (Coronado 2007). Second, it may be that the use of subjectivity to beneficially adjust measures for their insensitivity/incongruity was muted because supervisors did not have the authority to subjectively weight the multiple measures comprising the incentive contract, which is a form of subjectivity that might lead to congruity improvement (Hoppe and Moers

¹² I estimate Model 1 in Table 9 just to see the degree to which economic variables explain adjustments.

2007). Third, it may be that supervisors adjusted measures with the intention of providing benefits to managers or to themselves, but not necessarily to the organization.¹³ Fourth, my sensitivity and congruity constructs may be measured with too much error.

H3 predicted that supervisors' subjective adjustments to objective performance measures would be negatively (positively) related to supervisors' perception of performance measure precision (noise). The significant positive coefficient on *Noise* (0.471, p < .01) documents an increased probability that a subjective adjustment is made to objective performance measures as the level of noise in the measure increases, providing results consistent with H3.

H4 and H5 predicted that supervisors' subjective adjustments to objective performance measures would be negatively related to supervisors' perceptions of performance measure completeness and verifiability, respectively. Because the factor analysis reported in Chapter III revealed one underlying latent construct for completeness and verifiability, the coefficient on the variable *Completeness/Verifiability* in Table 9 is a joint test of both H4 and H5. The coefficient is significant and negative as predicted (-0.393, p < .01). Thus, the probability of supervisors adjusting performance measures decreases as the measures become more complete and verifiable. This result is consistent with H4 and H5. Note that I was not able to disentangle completeness and verifiability. Thus, the adjustments I document could be in response to a deficiency in completeness, verifiability, or both.

H6 predicted that supervisors' subjective adjustments to performance measures would be positively related to supervisors' perceptions of performance measure

¹³ This explanation seems plausible given supervisors' objective function (i.e., none of the measures in their pay plan evaluates them on the performance of the division or of the organization).

manipulability. Interestingly, Model 1 in Table 9 documents a significant *negative* effect of *manipulability* on the likelihood of adjustment (p < .01). However, in the full model, the coefficient on *Manipulability* is not significant (-0.114, p > .10). Thus, it appears that, after controlling for the effects of the performance measures, divisions, division performance, and the pre-adjusted level of the objective measures, the manipulability of a performance measure does not affect the likelihood of a subjective adjustment. An explanation is that when managers are aware that supervisors can "re-adjust" manager performance on a measure for manipulations, managers have less incentive to exert effort toward non-productive manipulations; and when measures have not been manipulated, supervisors don't have a reason to re-adjust. H6 is not supported.

H7 predicted that supervisors' subjective adjustments to objective performance measures would be positively related to perceptions of manager influence activities. The coefficient on *InfluenceActivity* is positive and significant (1.164, p < .01), indicating that adjustments to performance measures are more likely the more that managers engage in activities to try and favorably influence supervisors' evaluations of their performance.¹⁴ Results are thus consistent with H7 and document a cost of the use of subjectivity. H8 predicted that supervisors' subjective adjustments to objective performance measures would be positively related to perceptions of supervisor favoritism. The coefficient on *Favoritism* in Table 9 Model 2 is positive and significant (0.602, p < .01), indicating that supervisors who are reported to show favoritism are in fact more likely to subjectively adjust performance measures in an effort to reward some managers over others. Results are thus consistent with H8, and document an additional cost of the use of subjectivity.

¹⁴ Given supervisors' objective function (i.e., one of the objective performance measures in their pay plan evaluates their ability to develop the managers under their supervision), it seems plausible that manager influence activities would sway supervisors' evaluations.

It is noteworthy that the Planning (-0.669, p < .01) and Professional Development (-0.558, p < .01) performance measure indicators were significantly less likely to be adjusted than was the Program Management performance measure indicator (left out of the estimation in Table 9). This suggests that task complexity explains supervisors' subjective adjustments to performance measures over and above the effect that task complexity has on noise, completeness, and verifiability. Specifically, task complexity is normally thought to lead to noise and incompleteness in performance measures because task complexity requires managers to allocate their effort across more dimensions of performance to execute a successful outcome (Holmstrom and Milgrom 1991; Murphy and Oyer 2003; Coronado and Krishnan 2007). Program Management is a measure of arguably the most complex task that auditors perform because it tries to encapsulate more performance dimensions than the other measures. Accordingly, the significant coefficients on the performance measure indicator variables suggests that task complexity does more than just add to the noise and incompleteness of a measure.¹⁵

All eight division indicators left in the model were significant, indicating divisional differences in the use of subjective adjustments. The *level* variable was not significant (0.113, p > .10), indicating that the non-empirical negative association built into the scales did not impact supervisors' performance measure adjustments. Finally, *DivisionPerformance* was not significant (-0.094, p > .10).

¹⁵ An alternative explanation is that supervisors might have adjusted the Program Management measure more often because the organization did not allow supervisors to subjectively place more weight on this measure.

5.2. Sensitivity Tests

I conducted a variety of sensitivity tests, all resulting in qualitatively consistent results. For example, I included the downward adjustments omitted from the main analyses and estimated both a probit (Table A1) and an ordered probit model (Table A2).¹⁶ In the ordered probit model, the *Noise* factor explained less, consistent with the evidence of asymmetric adjustments for the effects of uncontrollable factors (Merchant 1989). I estimated probit models without including indicator variables for each of the objective performance measures (Table A3) and with a different variable for division performance (Table A4). I estimated probit models using an alternative definition of the pre-adjusted level of the objective measures (Table A5).¹⁷ I estimated probit models where I defined subjectivity as the presence of an adjustment that quantitatively changed the amount of the manager's incentive pay (Table A6). I ran the analyses using Huber-White cluster-corrected standard errors (Table A7). Finally, I estimated a probit model including the number of audit managers (NumberManagers) in the division as an additional control variable (Table A8). All of the empirical analyses provided qualitatively consistent results.

5.3. Supplementary Analysis and Discussion

Prior research shows that subjective performance evaluations often exhibit leniency and compression biases (Taylor and Wherry 1951; Murphy and Cleveland 1991). Leniency (compression) refers to skew (kurtosis) in the evaluation distribution. These biases are problems because performance evaluation systems are designed to

¹⁶ In the ordered probit model, subjective adjustments took on values of -1, 0, or 1.

¹⁷ Specifically, I measured the level variable as the difference between the pre-adjusted level of the objective measure and the population mean for that measure.

motivate optimal effort levels and to differentiate high performers (Bretz, Milkovich, and Read 1992). However, assuming performance is normally distributed, the presence of skewness and kurtosis suggest that managers may not have exerted maximum effort and/or that high performers have not been distinguished from low performers. For example, in the extreme case, compression is problematic because both poor-performing managers and superior-performing managers will be rated the same even though the superior-performing manager should have received a higher rating. In such a case, superior-performing managers have little motivation to work hard if they are going to be rated the same as poor-performing managers. Similarly, managers who perform poorly but receive lenient ratings have little motivation to exert more effort in the future. These managers do not likely even realize that they have performed poorly nor will they understand what performance level is really desired of them. Because resources are limited, managers who receive lenient ratings likely receive more of the pay raises and promotions they deserve while the more deserving, truly good-performing managers wind up receiving less than they really should. Therefore, results of compression and leniency include inappropriate reward distributions and impaired incentive contracting.¹⁸

As Figure 2, Panel A shows, the final performance evaluation distribution for the 2005 evaluation year (prior to the introduction of the new performance evaluation system) exhibited negative skewness (-3.7) and positive kurtosis (26.6), characteristic of leniency and compression. Note however, in Panel B of Figure 2, that the skewness (.3) and kurtosis (-0.2) are significantly mitigated after the implementation of the new performance evaluation system. Thus, because the final 2006 evaluations more closely

¹⁸ Biased evaluations create other costs as well because firms often use the performance evaluation system for numerous personnel decisions other than incentive compensation.

approximate a normal distribution, the introduction of the new pay-for-performance incentive plan itself provided benefits in the form of reducing two of the most common performance evaluation biases despite the predominance of one-side adjustments (i.e., 93 (7) percent of the adjustments were upward (downward)).

5.4. Conclusion

Overall, my study provides evidence that supervisors subjectively adjusted performance measures both to compensate for deficiencies in the measures as predicted by economic theory and due to social influences. In particular, I found support that subjective adjustments to performance measures were made for performance measures that were noisy, incomplete, and/or unverifiable; there was no support that subjective adjustments were made for measures that were not sensitive to agent actions, were not congruent with organizational objectives, and were susceptible to manipulation despite variation in these properties. In addition, performance measures that captured more of the task multidimensionality associated with complex jobs were adjusted more often than performance measures that captured less task multidimensionality. However, the use of subjectivity also induced costs because supervisors subjectively adjusted performance measures in response to influence activities of managers and their own personal preferences.¹⁹

¹⁹ Although it would be ideal to provide a quantitative measure of the overall cost/benefit realization, such a specification would be incomplete because not all of the relevant determinants of the subjective adjustments were identified and there will likely be spillover effects from some of the ones that were identified.

CHAPTER 6: SUMMARY

Internal audit work is complex and requires managers to allocate their effort across numerous job dimensions. In most professional work environments where jobs are complex and multi-dimensional, objective performance measures may be deficient in appropriately indicating a manager's actual contributions toward organizational objectives. In such environments, giving supervisors the authority to subjectively adjust the objective performance measures in a professional manager's incentive plan can reduce manager risk and align manager incentives with that of the organization. In other words, the use of subjective performance measure adjustments has the potential to compensate for the limitations of the objective performance measures. However, because of cognitive limitations and/or social biases of supervisors, the use of subjective performance measure adjustments can also introduce costs into incentive contracting.

I examined whether supervisors subjectively adjusted performance measures when they perceived there were deficiencies in the objective measures of managers under their supervision. I collected proprietary data on the performance evaluations of managers and survey data from three hierarchical management layers of a large internal audit organization. My tests reveal a number of interesting findings. First, supervisors with the authority to make upward or downward adjustments to objective performance measures make predominantly upward adjustments. Second, supervisors make these upward adjustments for performance measures they perceive to be incomplete, unverifiable, and noisy, but not for measures they perceive to be insensitive, incongruent, and susceptible to manipulation. Third, professional workers seem to hold the assumption that measures

37

that are sensitive, are by fiat congruent.²⁰ Fourth, supervisors are more (less) likely to subjectively adjust performance measures that try to capture more (less) of the task multidimensionality inherent in complex jobs. Fifth, the introduction of a pay-for-performance plan itself provided benefits in the form of eliminating two of the most common performance evaluation biases, leniency and compression. Finally, however, subjective performance measure adjustments introduced incentive contracting costs as supervisors subjectively adjusted performance measures in response to influence activities by managers and their own personal preferences.

This dissertation is subject to several limitations. First, I did not measure the "true" deficiencies in the performance measures. Instead, I asked survey participants to judge the deficiencies, and their subjective judgments likely involve some bias and/or noise. Nevertheless, their use of subjective adjustments should reflect their perceptions of performance measure deficiencies, regardless of whether those perceptions accurately portray the "true" values of the deficiencies. Second, consistent with prior subjectivity studies, my results suggest there are many yet-unidentified economic and psychological theoretical variables affecting subjective performance measure adjustments. Finally, the organization's new evaluation and compensation system represented a dramatically different way in which the performance of organizational members was evaluated. As a result of not being accustomed to the new system, it is possible that supervisors did not make adjustments in a "steady state" mode, but rather were still being influenced by the old system.

²⁰ Alternatively, highly sensitive measures may also be highly congruent measures, or my survey questions might not have distinguished performance measure sensitivity and congruity.

This dissertation has several implications for future research. The various uses of subjectivity are an increasingly important topic in performance evaluation and compensation systems (Ittner and Larcker 1998; MacLeod and Parent 1999). Subjectivity represents a potential for organizations to improve incentive contracting with managers when the objective performance measures used to evaluate managers fail to capture the multidimensional aspects of managers' complex jobs. There are clearly many factors that affect subjective performance measure adjustments, as well as the other uses of subjectivity, that future research should work to identify (Ittner et al. 2003).





Panel A – 2005 Final Evaluations (Prior to the introduction of the new performance evaluation system)

Panel B – 2006 Final Evaluations (Subsequent to the introduction of the new performance evaluation system)



Analysis of Non-response Bias

	N	leans
Criteria	Respondents	Non-respondents
# of shares distributed	2.26	2.41
	(p	= .52)
Incentive Payout	\$1,882.85	\$1,985.57
	(p	= .60)
% of managers with ≥ 1 adjustment	52.6 %	62.7 %
	(p	= .27)

Summary of Procedures to Arrive at Final Performance Measure Sample

Managers evaluated in the 2006 evaluation cycle	130
Deleted:	
Managers whose evaluation was "incomplete" as of July 27, 2007	7
Managers with "unique" performance measures	4
Duplicate evaluation record	1
FOIA data unmatched to agency-provided evaluation data	7
Total number of manager evaluation records	111
Number of performance measures in each manager's evaluation plan	<u>x 4</u>
Total number of performance measures available	444
Number of downwardly adjusted measures	<u>- 10</u>
Total number of performance measures in sample	<u>434</u>

Directorate	<u>Division</u>	<u>PM 1</u>	<u>PM 2</u>	<u>PM 3</u>	<u>PM 4</u>	<u>Total</u>
А	1	4	4	4	3	15
	2	1	4	4	4	13
	3	0	6	1	2	9
	4	<u>4</u>	<u>5</u>	<u>3</u>	<u>3</u>	<u>15</u>
	Dir A Totals	9	19	12	12	52
В	5	0	5	4	1	10
	6	0	5	3	1	9
	7	1	2	4	0	7
	8	<u>1</u>	<u>3</u>	<u>0</u>	<u>0</u>	<u>4</u>
	Dir B Totals	2	15	11	2	30
С	9	1	4	2	0	7
	10	1	1	2	3	7
	11	4	2	4	7	17
	12	<u>1</u>	<u>7</u>	<u>3</u>	<u>4</u>	<u>15</u>
	Dir C Totals	7	14	11	14	46
	Grand Total	18	48	34	28	128

Adjustments by Performance Measure, Division, and Directorate

PM 1 = Planning measure

PM 2 = Program Management measure

PM 3 = Reporting measure

PM 4 = Professional Development measure

	Table 4 Common Factor Analysis with Varimax Rotation of the 12 P. (All factor loadings greater than	Performance Meas .3 are shown, n = 1	ure Characteristic 12)	c Survey	Items
Ite	Suc	Sensitivity/ Congruity	Verifiability/ Completeness	Noise	Manipulability
ب م ا	The (organization's) missions and goals are further				
	accomplished when AMs perform well on this (performance measure)	0.86			
6.	If AMs perform well, it is directly reflected in better				
	performance on this (performance measure)	0.76			
ပ	When AMs perform well on this (performance measure), they contribute				
	to the (organization's) mission and goals	0.92			
ъ.	AM effort leads to better performance on this (performance measure)	0.69			
e.	Many of the important activities that AMs perform increase				
	performance on this (performance measure) but do not contribute to				
	the organization's mission and goals ^a	0.57			0.42
÷	This performance measure does not capture important Performance				
	measure name activities that AMs perform ^a		0.60		0.36
oio	This (performance measure) captures all the dimensions of effort that				
	are required for AMs in this division to perform well on this				
	(performance measure)		0.74		
h.	AM performance on this (performance measure) can be verified		0.74		
. _:	This (performance measure) is quantifiable		0.57		
·	AM performance on this (performance measure) is affected by				
	unanticipated events/changes			0.80	
¥.	AM performance on this (performance measure) is influenced				
	by things outside of their control			06.0	
	AMs can manipulate this (performance measure) to ensure they meet				
	performance goals without actually performing well				-0.89
Ei	gnenvalues	3.16	2.04	1.54	1.27
	These items were reverse-coded. Mc= Audit Managers				
2					

.

45

 Table 5

 Descriptive statistics of the 12 Performance Measure Characteristic Survey Items

Sel .	nsitivity/Congruity The (orcenization's) missions and acals are further accountished when	<u>Mean</u>	Std. Dev.	Range	Z
v	AMS perform well on this (performance measure)	5.83	1.33	1 – 7	112
6.	If AMs perform well, it is directly reflected in better performance on				
	this (performance measure)	5.38	1.49	1 – 7	112
ပ	When AMs perform well on this (performance measure), they contribute to the				
	(organization's) mission and goals	5.79	1.33	1 – 7	112
Ч.	AM effort leads to better performance on this (performance measure)	6.04	1.11	1 – 7	112
e.	Many of the important activities that AMs perform increase performance on				
	this (performance measure) but do not contribute to the organization's mission				
	and goals ^a	4.84	1.67	1 – 7	111
Ve	rifiability/Completeness				
ч <u>.</u>	This performance measure does not capture important Performance measure name				
	activities that AMs perform ^a	4.53	1.96	1 – 7	111
ò	This (performance measure) captures all the dimensions of effort that are				
	required for AMs in this division to perform well on this (performance measure)	4.13	1.93	1 – 7	112
ų.	AM performance on this (performance measure) can be verified	4.66	1.86	1 – 7	111
. 	This (performance measure) is quantifiable	5.68	1.32	1 – 7	111
ž	Dise				
·	AM performance on this (performance measure) is affected by				
	unanticipated events/changes	4.58	1.91	1 – 7	112
ĸ.	AM performance on this (performance measure) is influenced				
	by things outside of their control	4.53	1.77	1 - 7	112
Σ	anipulability				
ij.	AMs can manipulate this (performance measure) to ensure they meet				
	performance goals without actually performing well	3.38	1.76	1 – 7	111

^a These items were reverse-coded.

	Measure
	Performance
	Å
able 6	Factors
	4
	of the
	statistics (
	escriptive
	Ă

		PM 1			PM 2			PM 3			PM 4	
Factor	Mean	St Dev	Range	Mean	St Dev	Range	Mean	St Dev	Range	Mean	<u>St Dev</u>	Range
Sensitivity/	0.25	0.45	[-0.79, 0.88]	0.56	0.34	[01.1,00.0-]	0.56	0.27	[0.06, 0.94]	-0.72	0.71	[-1.70, 0.90]
Congruity												
Completeness/	0.07	0.42	[-0.65, 0.69]	0.01	66.0	[-1.85, 1.21]	-0.18	0.89	[-1.61, 0.69]	0.17	0.78	[-1.02, 1.24]
Verifiability												
Noise	-0.28	0.80	[-1.56, 0.54]	0.45	0.54	[-1.11, 1.06]	-0.04	0.57	[-0.77, 0.97]	0.19	0.79	[-1.10, 1.41]
Manipulability	-0.46	0.90	[-1.59, 0.93]	0.01	0.58	[-1.31, 0.70]	0.22	0.53	[-0.76, 1.08]	-0.24	0.66	[-1.38, 1.00]

PM 1 = Planning measure PM 2 = Program Management measure PM 3 = Reporting measure PM 4 = Professional Development measure

Descriptive Statistics of Costs of Subjectivity and Control Variables

(n = 434)

<u>Variable</u>	Mean	Std. Dev.	Actual <u>Range</u>
Influence Activity	4.50	0.73	3.5 - 6.0
Favoritism	2.82	0.79	1.3 - 4.0
Division Performance	16.97	0.92	15.00 - 18.11
Level	3.31	0.60	2 - 5

 Table 8

 Correlation Matrix of the Independent Variables

	<u>Variable</u> Sensitivity/Congruity Completeness/Verifiability Noise Manipulability Influence Activity Favoritism Planning Program Management Reporting Professional Develonment	1 -0.13 -0.15 -0.06 -0.09 0.09 0.35 0.35	2 -0.10 0.20 0.27 0.27 0.04 0.04 0.04 0.04 0.14	<u>3</u> -0.09 -0.13 -0.13 -0.13 -0.28 -0.10	4 0.14 0.21 0.21 0.28 0.27 0.10 0.27	5 0.18 0.18 0.09 0.00 0.00 0.00 0.00 0.00 0.00	6 -0.02 -0.20 -0.10 0.10 0.00 0.00 0.00	7 -0.03 -0.07 -0.21 -0.01 -0.01 -0.33	8 0.27 0.05 0.35 0.09 0.00 0.00 0.33 0.33	9 0.35 0.11 0.13 0.28 0.28 0.28 0.33 0.33	10 0.58 0.15 0.05 0.05 0.01 0.00 0.01 0.33	11 0.04 0.01 0.29 0.00 0.00 0.00 0.00	11 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0
11.	Division Performance	0.04	-0.03	-0.15	0.00	0.35	0.08	0.00	0.01	0.00	0.00		0.04
.71	revel	~~~~	r	10.0	0.00	71.0-	-0.0-	CO.0-	71.0	0.00	N.)-	3.0	

Pearson (Spearman) correlation coefficients are presented below (above) the diagonal. Bolded correlations are significant at the 10% level (two-tailed).

	Pred		Model	1		Model	2
	<u>Sign</u>	Coeff.	z-stat	$\frac{dy}{dx}^{c}$	Coeff.	<u>z-stat</u>	$\frac{dy}{dx}^{c}$
Intercept		-0.554	-8.17***		-5.690	-2.40**	
SensitivityCongruity	-	0.073	0.77	0.02	-0.099	-0.51	-0.03
Noise	+	0.236	2.52***	-0.06	0.471	2.67***	0.15
CompletenessVerifiability	-	-0.163	-2.02**	0.08	-0.393	-2.56***	-0.13
Manipulability	+	-0.261	-2.66	0.09	-0.114	-0.70	-0.04
InfluenceActivity	+				1.164	2.82***	0.38
Favoritism	+				0.602	3.05***	0.20
Planning					-0.669	-2.89***	-0.19
Reporting					-0.214	-1.07	-0.07
ProfessionalDevelopment					-0.558	-1.86*	-0.17
Div ₂ ^b					1.437	2.98***	0.53
Div ₄					-2.300	-2.47**	-0.34
Div ₈					0.787	1.91*	0.29
Div9					-2.307	-3.01***	-0.32
Div ₁₀					1.500	3.12***	0.55
Div ₁₁					-2.187	-3.28***	-0.33
Div ₁₂					0.184	0.46	0.06
DivisionPerformance					-0.094	-0.98	-0.03
Level					0.113	0.98	0.04
Pseudo \mathbf{R}^2			0.029			0 109	
% of Correct Classifications			60 50%			72 120	
Sample size)		121			12.12.10	
Sample Size			434			434	

Probit Regression of Subjective Performance Measure Adjustments^a

^a Dependent variable = 1 (0) if performance measure was (was not) adjusted (A).

^b Div_i are the division indicators. I dropped five division indicators due to multicollinearity.

^c The dy/dx columns report the marginal effects of the independent variables on the probability of adjustment.

***, **, * Statistically significant at the 1%, 5%, and 10% levels, respectively (one-tailed for the predicted coefficients and two-tailed for the non-predicted coefficients).

APPENDICES

Appendix A Review of the Literature on Subjectivity in Performance Evaluation and Compensation

Subjective signal studies

		And and a second s	A CONTRACTOR OF A CONTRACTOR O			
Study	Consideration that	Hierarchical	Nature of signal	Private	Control	Key Result
	supervisor's	Structure	that supervisor	Information	Svstem	
	incentive may		receives about	Structure	Examined?	
	conflict with that		manager's	Between		
	of the principal?		performance	Contracting		
				Parties		
Baiman and	No – supervisor is	One	Verifiable and	1. Jointly	Bonus Pool	Using the noncontractible
Rajan 1995	residual claimant.	supervisor;	non-verifiable	observable and	Schemes	signal to pay discretionary
		two	signals.	verifiable.		bonus awards improves
		managers.		2. Only		incentive contracting
				supervisor		compared to only using the
				observes non-		contractible signal because it
				contractible		improves risk sharing
				signal.		between the two managers.
Baker et al.	No - supervisor is	One	Verifiable and	Both signals	N/A	When subjective signal is
1994 (also	residual claimant.	supervisor,	non-verifiable	are jointly		used to complement objective
look at		one manager.	signals.	observable and		manipulable signal, incentive
weighting)				verifiable.		contracting is improved.

																	-		-						_
	Key Result						Discretionary bonus pools	are optimal when	supervisors rely on	subjective signals. But it	also imposes additional risk	on managers; the risk	decreases as the subjective	signal becomes more	precise or as the number of	managers in the bonus pool	increases.	Allowing supervisors to use	their private signal of	manager performance	improves incentive	contracting.			
	Control	System	Examined?				Bonus pools											Bonus pool	size and	allocation.					
	Private	Information	Structure	Between	Contracting	Parties	Privately	observed by	supervisor.									Supervisor and	managers	observe group	output signal;	only supervisor	observes each	manager's	output.
	Nature of signal	that supervisor	receives about	manager's	performance		Verifiable and	non-verifiable	signals.									1. Group	output.	2. Each	manager's	output.			
	Hierarchical	Structure					One	supervisor,	"n"	managers.								One	supervisor;	two	managers.				
(maning)	Consideration	that supervisor's	incentive may	conflict with that	of the principal?		No - supervisor	is residual	claimant.									No – supervisor	is residual	claimant.					
on a vininday	Study						Rajan and	Reichelstein	2007 (also	look at	weighting)							Fisher et al.	2005						

-	
(par	
atint	
(cor	
<	
Ľ,	
pu	
പ	
8	

ure of sig supervisis supervision <u>ager's</u> <u>ormance</u> ervisor rracted reaction all.	bon estimation of the state of	Hierarchical Mat Structure thatian Interaction trescondance One performance Supervisor, condance Number of the structure info Managers, info Supervisor 1. E and manager, condance Supervisor 1. E and manager, condance of the of the structure of the structure		ure of signal Private Control System Key Result	supervisor Information Examined?	lager S between	ormance Contracting	Parties	ervisor Supervisor Bonus plan Bonus plan type influences	ives a non- knows each type (partial or supervisors' use of non-	tracted manager's full discretion contracted information.	rmation signals. for allocating Supervisors with partial	al. bonus is given discretion do not fully	to managers). incorporate non-contracted	information.	xplicit N/A N/A Subjective bonus payouts are	tract associated with departmental	interdependencies, financial	ubjective losses, and achievability of	als outside formula bonuses and affect pay	xplicit satisfaction, productivity, and	us award profitability (conditional on	
A Consideration that Consideration that incentive may conflict with that incentive may conflict with that incentive may conflict with that incention and the principal? First and First and Fir	A (continued) Consideration that incentive may conflict with that of the principal? Yes – supervisor has a cognitic visit (uses an anchor has a a and ainst (uses an anchor has a a and ainst (uses an anchor has a cognitic to the principal has a cognitic has a cognitic has a cognitic to the principal has a cognitic to the principal has a cognitic has a cognit		Appendix	Study					Bailey	et al.	2006					Gibbs	et al.	2004					

ontinue
A (c
vipuec
ď

	Key Result						There are two	explanations for the	observed positive relation	between risk and	incentives: (1) Supervisor	bias (favoritism) in	subjective evaluations	and (2) costly	investigations to obtain	an accurate measure of	manager output.	Because of supervisor	bias (favoritism), the firm	puts too little weight on	subjective evaluations	and excessive emphasis	on bureaucratic rules	which means they put	excess weight on the	verifiable signal.
	Control System	Examined?					Costly	investigations	by the	supervisor of	the signal.							The firm	monitors the	supervisor.						
	Private	Information	Structure	Between	Contracting	Parties	Supervisor	privately observes	the noisy non-	verifiable signal	and knows with	certainty his/her	preference for	favoritism.				Supervisor	privately observes	the noisy non-	verifiable signal	and knows with	certainty his/her	preference for	favoritism.	
	Nature of	signal that	supervisor	receives about	manager's	performance	Verifiable and	non-verifiable	signals.									Verifiable and	non-verifiable	signals.						
	Hierarchical	Structure					The firm, one	supervisor,	and one	manager.								Firm, one	supervisor,	one manager.						
aluation studies	Is there	consideration	that supervisor's	incentive may	conflict with that	of the principal?	Yes - favoritism.											Yes - favoritism.								
Subjective ev	Study						Prendergast	2002										Prendergast	and Topel	1996						

Appendix A (continued)

al a summaday						
Study	Is there	Hierarchical	Nature of signal	Private	Control	Key Result
	consideration that	Structure	that supervisor	Information	System	
	supervisor's		receives about	Structure	Examined?	
	incentive may		manager's	Between		
	conflict with that		performance	Contracting		
	of the principal?			Parties		
MacLeod	Yes -	One	Subjective signal.	Both	Supervisor	The quality of the subjective
2003	discrimination.	supervisor,		supervisor	can divert	signal is not often as
		one		and manager	payout to	important as the correlation
		manager.		observe	third part if	between the supervisor and
				signals.	he/she thinks	manager's beliefs about that
					the manager	signal. When supervisor's
					is shirking.	and manager's beliefs about
						the subjective signal are not
						correlated, the optimal
						contract results in more
						compressed pay due to
						"conflict" costs between
						supervisor and manager.
Hopwood	No	Supervisor	Managers are	Supervisor	N/A	Job tension, poor relations,
1972		and	evaluated on	privately		and manipulation of
		manager.	objective	observes non-		objective measures is more
			measures (budget	verifiable		likely the more that
			constrained) on	signal.		managers perceive they are
			or supervisor's			evaluated on objective
			subjective			measures only (as opposed
			judgment) (non-			to being subjectively
			accounting style)			evaluated).
Duckeye at	N		11-17-11		×11.4	
busnman et	NO.	Supervisor	Verifiable and	Supervisor	N/A	Growth opportunities and

ppendix A (continue

length of product time horizon (but not the noise in verifiable signals) all increase the use of	evaluations. The effectiveness of any subjective performance evaluations depends on business unit strategy. It is with a "build" strategy but hampers it in business units	The overall number of must a number of measures in the explicit measures in the explicit subjective measures used to evaluate managers are positively related to proformance evaluation compression and lenterey.	Environmental uncertainty increases the use of subjective performance evaluations.
	N/A	N/A	N/A
privately observes non- verifiable signal.	Supervisor privately observes non- verifiable signal.	Supervisor privately observes non- verifiable signal.	Supervisor privately observes non- verifiable signal.
non-verifiable signals.	Verifiable and non-verifiable signals.	Verifiable and non-verifiable.	Verifiable and non-verifiable signals.
and manager.	Supervisor and manager.	Supervisor and manager.	Supervisor and manager.
	No.	Yes – supervisors are prone to compression and leniency biases.	No
al. 1996	Govindarajan and Gupta 1985	Moers 2005	Govindarajan 1984

Subjective weighting studies

Key Result	Supervisors weighted financial measures more than non- financial measures, incorporated factors other than socrecard tactors other than socrecard enanged evaluation criteria from quarter to quarter, did no quarter of quarter, did no predictive of future financial measures that were not measures that were not performance.	Although individuals are hittoread by changes in the precision and error covariance of cost and quality measures when they make subjective derisions, they orden do not make adjustments in the same direction predicted by agency theory.
Control System Examined?	Top management reviews of supervisor evaluations.	N/A
Private Information Structure Between Contracting Parties	Supervisor phivately observes non- verifiable signal.	N/A
Nature of signal that supervisor receives about manager's performance	Verifiable and non-verifiable signals.	Verifiable signals.
<u>Hierarchical</u> <u>Structure</u>	Supervisor and manager	Supervisor and manager.
Is there consideration that consideration that supervisor's incentive may conflict with that of the principal?	Yes – supervisors exhibited a number of biases.	Yes - not all individuals engage in the long chain of reasoning required by agency theory.
Study	littner, Larcker, and Meyer 2003	Krishnan, Luft, and Shields 2005

	<u>Key Result</u>	Individuals weight measures that are common across divisions more in evaluating division performance than they weight measures that are unique across divisions.
	Control System Examined?	N/A
	Private Information Structure Between Contracting Parties	N/A
	Nature of signal that supervisor receives about manager's performance	Verifiable and non-verifiable.
	Hierarchical Structure	Supervisor and manager.
(nontrino) a	<u>Is there</u> consideration that supervisor's incentive may conflict with that of the principal?	Yes – supervisors have cognitive biases.
T VINIA day	Study	Lipe and Salterio 2000

Appendix A (continued)

Appendix B: Survey Instrument

1. Survey questions of the properties of performance measures. Answering scale was 1 - 7 (Strongly Disagree – Strongly Agree). Means (standard deviations) are shown by division manager (DM) and supervisor (S). Performance measures were randomized, but the questions for each measure were not. Survey responses were obtained for these questions from 8 DMs and 20 Ss, for a total of 28 responses.

To what extent do you agree or disagree with the following statements about the *PLANNING* performance measure:

Congruity	DM	<u>S</u>
a. The organization's missions and goals are further accomplished when	6.1	5.7
audit managers (AMs) perform well on this performance measure	(0.8)	(1.4)
b. When AMs perform well on this performance measure, they contribute	6.0	5.9
to the organization's mission and goals	(0.8)	(1.3)
c. Many of the important activities that AMs perform increase performance	2.7	3.4
on this performance measure but do not contribute to the organization's	(1.3)	(1.9)
mission and goals		
Sensitivity		
d. If AMs perform well, it is directly reflected in better performance on this	5.5	5.1
performance measure	(1.2)	(1.6)
e. AM effort leads to better performance on this performance measure	6.3	5.9
	(0.7)	(1.5)
f. This performance measure does not capture important Planning activities	3.2	3.6
that AMs perform	(1.5)	(2.0)
Precision		
g. AM performance on this performance measure is affected by	3.1	4.1
unanticipated events/changes	(1.8)	(2.1)
h. AM performance on this performance measure is influenced by things	3.3	4.2
outside of their control	(1.9)	(2.1)
Completeness		
i. This performance measure captures all the dimensions of effort that are	4.8	3.7
required for AMs in this division to perform well on this performance	(1.9)	(1.9)
measure		
Verifiability		
j. This performance measure is quantifiable	5.4	5.4
	(1.6)	(1.5)
k. AM performance on this performance measure can be verified	6.3	6.1
	(0.7)	(0.9)
<u>Manipulability</u>		
1. AMs can manipulate this performance measure to ensure they meet	3.6	4.1
performance goals without actually performing well	(2.6)	(1.7)

Appendix B (continued)

Quality/Quantity	DM	<u>S</u>
m. This performance measure focuses on the quality of AM performance	5.3	3.7
	(1.0)	(2.3)
n. This performance measure focuses on the quantity of AM performance	5.6	5.6
	(0.7)	(2.0)

To what extent do you agree or disagree with the following statements about the **PROJECT MANAGEMENT** performance measure:

Congruity	DM	<u>S</u>
a. The organization's missions and goals are further accomplished when AMs	6.1	6.4
perform well on this performance measure	(0.8)	(1.0)
b. When AMs perform well on this performance measure, they contribute to	6.4	6.4
the organization's mission and goals	(0.9)	(0.9)
c. Many of the important activities that AMs perform increase performance on	2.7	2.7
this performance measure but do not contribute to the organization's mission	(1.7)	(1.4)
and goals		
Sensitivity		
d. If AMs perform well, it is directly reflected in better performance on this	6.3	6.2
performance measure	(0.9)	(0.7)
e. AM effort leads to better performance on this performance measure	6.6	6.5
	(0.7)	(0.6)
f. This performance measure does not capture important Project Management	3.1	3.3
activities that AMs perform	(2.0)	(2.0)
Precision		
g. AM performance on this performance measure is affected by unanticipated	4.9	5.7
events/changes	(1.2)	(1.5)
h. AM performance on this performance measure is influenced by things	4.6	5.5
outside of their control	(1.1)	(1.3)
Completeness		
i. This performance measure captures all the dimensions of effort that are	4.9	4.5
required for AMs in this division to perform well on this performance measure	(1.8)	(2.2)
Verifiability		
j. This performance measure is quantifiable	4.8	4.1
	(1.7)	(2.0)
k. AM performance on this performance measure can be verified	5.8	5.6
	(0.7)	(1.6)
<u>Manipulability</u>		
1. AMs can manipulate this performance measure to ensure they meet	2.8	3.3
performance goals without actually performing well	(1.4)	(1.5)
Quality/Quantity		
m. This performance measure focuses on the quality of AM performance	6.0	5.4
	(1.1)	(1.8)
n. This performance measure focuses on the quantity of AM performance	4.9	4.8
	(1.5)	(1.9)

Appendix B (continued) To what extent do you agree or disagree with the following statements about the **REPORTING** performance measure:

Congruity	DM	<u>S</u>
The organization's missions and goals are further accomplished when AMs	6.6	6.2
perform well on this performance measure	(0.5)	(1.0)
When AMs perform well on this performance measure, they contribute to	6.8	6.0
the organization's mission and goals	(0.5)	(1.5)
Many of the important activities that AMs perform increase performance	2.2	2.6
on this performance measure but do not contribute to the organization's	(1.3)	(1.0)
mission and goals		
Sensitivity		
If AMs perform well, it is directly reflected in better performance on this	6.5	5.8
performance measure	(0.8)	(1.4)
AM effort leads to better performance on this performance measure	6.6	6.1
	(0.5)	(1.0)
This performance measure does not capture important Reporting activities	2.5	3.3
that AMs perform	(1.6)	(2.1)
Precision		
AM performance on this performance measure is affected by unanticipated	4.3	4.6
events/changes	(1.6)	(2.1)
AM performance on this performance measure is influenced by things	4.0	4.5
outside of their control	(1.4)	(1.9)
Completeness		
This performance measure captures all the dimensions of effort that are	5.1	3.9
required for AMs in this division to perform well on this performance	(1.6)	(2.1)
measure		
Verifiability		
This performance measure is quantifiable	4.6	4.1
	(1.2)	(2.3)
AM performance on this performance measure can be verified	6.0	5.4
	(0.8)	(1.6)
<u>Manipulability</u>		
AMs can manipulate this performance measure to ensure they meet	2.1	2.9
performance goals without actually performing well	(1.1)	(1.3)
Quality/Quantity		
This performance measure focuses on the quality of AM performance	6.1	5.1
	(0.6)	(1.8)
This performance measure focuses on the quantity of AM performance	4.1	4.1
	(1.6)	(2.2)
Appendix B (continued)

To what extent do you agree or disagree with the following statements about the **PROFESSIONAL DEVELOPMENT** performance measure:

Congruity	DM	S
The organization's missions and goals are further accomplished when AMs	4.8	5.0
perform well on this performance measure	(1.8)	(1.5)
When AMs perform well on this performance measure, they contribute to	4.6	4.7
the organization's mission and goals	(1.4)	(1.3)
Many of the important activities that AMs perform increase performance	3.9	3.7
on this performance measure but do not contribute to the organization's	(1.9)	(1.8)
mission and goals		
Sensitivity		
If AMs perform well, it is directly reflected in better performance on this	4.3	4.1
performance measure	(1.7)	(1.4)
AM effort leads to better performance on this performance measure	5.5	5.4
	(0.9)	(1.3)
This performance measure does not capture important Professional	3.9	3.2
Development activities that AMs perform	(2.0)	(2.0)
Precision		
AM performance on this performance measure is affected by unanticipated	4.3	4.7
events/changes	(2.1)	(1.9)
AM performance on this performance measure is influenced by things	4.3	4.8
outside of their control	(1.6)	(1.7)
Completeness		
This performance measure captures all the dimensions of effort that are	3.8	3.8
required for AMs in this division to perform well on this performance	(1.8)	(1.7)
measure		
Verifiability		
This performance measure is quantifiable	4.9	4.7
	(2.0)	(1.9)
AM performance on this performance measure can be verified	6.0	5.2
	(0.8)	(1.6)
Manipulability		
AMs can manipulate this performance measure to ensure they meet	4.7	3.6
performance goals without actually performing well	(2.0)	(2.0)
Quality/Quantity		
This performance measure focuses on the quality of AM performance	4.3	3.5
	(1.8)	(1.9)
This performance measure focuses on the quantity of AM performance	5.4	4.5
	(0.7)	(2.1)

Appendix B (continued)

2. Survey questions of influence activities. Answering scale was 1 - 7 (Very Little – A Great Deal). Means (standard deviations) are shown by division manager (DM), supervisor (S), and audit manager (AM). Survey responses were obtained for these questions from 8 DMs, 20 Ss, and 40 AMs, for a total of 68 responses.

Items

Please rate the following questions about the job of an AM:

	DM	<u>S</u>	<u>AM</u>
a. (1) How much time do you have to spend trying to favorably	N/A	N/A	3.1
influence your PM's assessment of your performance? ¹			(2.0)
a. (2) To what extent do AMs try to favorably influence your	4.6	4.6	N/A
assessment of their performance? ²	(1.1)	(2.1)	
b. (1) How much time do you have to spend trying to favorably	N/A	N/A	2.9
influence the impression that your PM has of you? ¹			(1.8)
b. (2) To what extent do AMs try to favorably influence your	2.4	4.7	N/A
impression of them? ²	(1.3)	(1.8)	

¹ This question was asked to AMs.

² This question was asked to Ss and DMs.

3. Survey questions of favoritism. Answering scale was 1 - 7 (Strongly Disagree – Strongly Agree). Means (standard deviations) are shown by supervisor (S) and audit manager (AM). Survey responses were obtained for these questions from 20 Ss and 19 AMs, for a total of 39 responses.

Items

Please indicate the extent to which you agree or disagree with each of the following statements about the way you (your supervisor) adjusted performance measures in the performance evaluation cycle just completed. I (My supervisor) used adjustments:

	<u>S</u>	<u>AM</u>
a. (1) To improve the quality of his/her personal relationship with me ¹	N/A	2.4
a. (2) To improve the quality of my personal relationship with some AMs ²		(1.4)
	1.5	N/A
	(1.1)	
b. (1) To increase rewards for some AMs relative to other AMs under	N/A	3.8
his/her supervision ¹		(1.8)
b. (2) To increase rewards for some AMs relative to other AMs under my	3.9	N/A
supervision ²	(2.1)	

¹ This question was asked to AMs.

² This question was asked to Ss.

Probit Model Including Downward Adjustments^a

	Pred	Model	
	<u>Sign</u>	Coeff.	<u>z-stat</u>
Intercept		13.102	1.81
SensitivityCongruity	-	-0.040	-0.22
Noise	+	0.295	1.71**
CompletenessVerifiability	-	-0.179	-1.45*
Manipulability	+	-0.084	-0.49
InfluenceActivity	+	0.813	2.09**
Favoritism	+	0.181	1.58*
Planning		-0.663	-2.70***
Reporting		-0.297	-1.46*
ProfessionalDevelopment		-0.602	-1.96**
Div ₂ ^b		0.751	2.14**
Div ₄		-0.516	-1.01
Div ₅		-0.742	-2.08**
Div ₈		-0.367	-0.68
Div ₉		-0.854	-1.24
Div ₁₀		1.292	2.39**
Div ₁₁		-1.970	-2.39**
Div ₁₂		0.164	0.42
DivisionPerformance		-0.324	-2.73***

Pseudo R^2	0.107
% of Correct Classifications	71.85%
Sample size	444

^a Dependent variable = 1 (0) if performance measure was (was not) adjusted (A).

^b Div_i are the division indicators.

Ordered Probit Model Including Downward Adjustments^a

	Pred	<u>N</u>	<u>lodel</u>
	<u>Sign</u>	Coeff.	<u>z-stat</u>
SensitivityCongruity	-	-0.092	-0.50
Noise	+	0.227	1.50*
CompletenessVerifiability	-	-0.233	-1.81**
Manipulability	+	-0.010	-0.07
InfluenceActivity	+	9.158	1.84**
Favoritism	+	3.562	1.82**
Planning		-0.503	-2.42**
Reporting		-0.154	-0.82
ProfessionalDevelopment		-0.333	-1.18
Div ₂ ^b		9.731	1.82*
Div ₄		-16.110	-1.84*
Div ₅		7.763	1.71*
Div ₈		0.782	2.38**
Divo		-10.164	-1.94*
Div ₁₀		6.684	1.95*
Div ₁₁		-8.140	-1.95*
Div ₁₂		2.569	1.56
DivisionPerformance		-0.503	-2.40**
Pseudo \mathbf{R}^2		0.0	75
Sample size	0.075		
Sample Size		444	

^a Dependent variable = 1 or -1 (0) if performance measure was (was not) adjusted (A)

upward or downward, respectively.

^b Div_{i} are the division indicators.

Probit Model without Performance Measure Indicators^a

	Pred	<u>Model</u>	
	<u>Sign</u>	Coeff.	<u>z-stat</u>
Intercept		-3.650	1.43
SensitivityCongruity	-	0.113	0.97
Noise	+	0.403	2.97***
CompletenessVerifiability	-	-0.443	-3.04***
Manipulability	+	0.084	0.60
InfluenceActivity	+	1.330	3.32***
Favoritism	+	0.506	2.54***
Div ₂ ^b		1.558	3.28***
Div ₄		-2.501	-2.73***
Div ₅		-0.556	1.88*
Div ₈		-1.771	-2.26**
Divo		1.713	3.72***
Div ₁₀		-1.963	-2.88***
Div ₁₁		0.106	0.29
DivisionPerformance		-0.279	-2.65***
Level		0.180	1.55
$\mathbf{P}_{saudo} \mathbf{P}^2$		0.0	166
af Compat Classifications		71	000 6601
% of Correct Classifications		/1	.00%
Sample size		434	

^a Dependent variable = 1 (0) if performance measure was (was not) adjusted (A).

^b Div_i are the division indicators.

Probit Model with Alternative Division Performance Variable^a

	Pred	Model	
	<u>Sign</u>	Coeff.	<u>z-stat</u>
Intercept		-5.997	-2.63***
SensitivityCongruity	-	-0.097	-0.49
Noise	+	0.408	2.26**
CompletenessVerifiability	-	-0.433	-2.84***
Manipulability	+	-0.103	-0.63
InfluenceActivity	+	1.042	2.58***
Favoritism	+	0.723	2.74***
Planning		-0.692	-3.02***
Reporting		-0.248	-1.23
ProfessionalDevelopment		-0.566	-1.86
Div ₂ ^b		1.084	2.24**
Div ₄		-2.192	-2.29**
Div ₅		0.455	1.49
Div ₈		-2.161	-2.62***
Div ₉		0.886	1.86*
Div ₁₀		-1.535	-1.93**
Div ₁₁		0.103	0.25
DivisionPerformance		-0.256	-1.71*
Level		0.098	0.86
$P_{seudo} P^2$		Δ	103
The of Correct Classifications		0. 70	Q102
Somple size		/2.81%	
Sample size		43	4

^a Dependent variable = 1 (0) if performance measure was (was not) adjusted (A).

^b Div_i are the division indicators.

Probit Model with Alternative Level Variable^a

	Pred	<u>Model</u>	
	<u>Sign</u>	Coeff.	<u>z-stat</u>
Intercept		-4.277	-1.59
SensitivityCongruity	-	-0.074	-0.37
Noise	+	0.242	1.57*
CompletenessVerifiability	-	-0.433	-2.85***
Manipulability	+	0.002	0.01
InfluenceActivity	+	1.447	3.39***
Favoritism	+	0.536	2.60***
Planning		-0.810	-3.49***
Reporting		-0.373	-1.91*
ProfessionalDevelopment		-0.594	-1.96**
Div ₂ ^b		1.846	3.68**
Div ₄		-2.399	-2.47**
Div ₅		0.726	2.34
Div ₈		-1.764	-2.10**
Div ₉		1.790	3.64***
Div ₁₀		-2.139	-2.98***
Div ₁₁		0.165	0.43
DivisionPerformance		-0.342	-3.13***
Level		0.627	3.92***
$\mathbf{P}_{\mathrm{regula}} \mathbf{P}^2$		0	127
		0.	13/
% of Correct Classifications		/3	0.90%
Sample size		434	

^a Dependent variable = 1 (0) if performance measure was (was not) adjusted (A).

^b Div_{i} are the division indicators.

Probit Model with Alternative Adjustment Variable^a

	Pred	Model	
	<u>Sign</u>	Coeff.	<u>z-stat</u>
Intercept		-4.053	-1.29
SensitivityCongruity	-	0.143	0.62
Noise	+	0.232	1.43*
CompletenessVerifiability	-	-0.405	-2.43***
Manipulability	+	0.097	0.54
InfluenceActivity	+	1.310	2.65***
Favoritism	+	0.620	2.62***
Planning		-0.069	-0.27
Reporting		-0.002	-0.01
ProfessionalDevelopment		-0.008	-0.02
Div ₂ ^b		1.261	2.15**
Div ₄		-2.132	-1.89*
Div ₅		0.082	0.22
Div ₈		-1.264	-1.34
Div9		1.621	2.94***
Div ₁₀		-1.842	-2.28**
DivisionPerformance		-0.319	-2.66***
Level		0.302	2.28**
$\mathbf{P}_{\text{seudo}} \mathbf{P}^2$		0	111
r seudo K		0.	111
% of Correct Classifications		82	
Sample size		434	

^a Dependent variable = 1 if the performance measure was adjusted (A) and changed the amount of the manager's incentive pay, or = 0 if the performance measure was adjusted but did not change the amount of the manager's incentive pay or if the performance measure was not adjusted.

^b Div_i are the division indicators.

Probit Model using Huber-White Cluster-Corrected Standard Errors^a

	Pred	<u>Model</u>		
	<u>Sign</u>	Coeff.	<u>z-stat</u>	
Intercept		-2.515	-0.96	
SensitivityCongruity	-	-0.056	-0.28	
Noise	+	0.222	1.28*	
CompletenessVerifiability	-	-0.432	-2.96***	
Manipulability	+	-0.001	-0.01	
InfluenceActivity	+	1.202	2.93***	
Favoritism	+	0.458	2.24**	
Planning		-0.762	-3.28***	
Reporting		-0.356	-1.80*	
ProfessionalDevelopment		-0.536	-1.81*	
Div ₂ ^b		1.512	3.15***	
Div ₄		-2.169	-2.43**	
Div ₅		0.577	1.83*	
Div9		-1.698	-2.07**	
Div ₁₀		1.437	3.05***	
Div ₁₁		-1.941	-2.71***	
Div ₁₂		0.038	0.10	
DivisionPerformance		-0.272	-2.36**	
Level		0.143	1.24	
$\mathbf{P}_{\mathrm{saudo}} \mathbf{P}^2$		0	110	
r scuud K		0. 70	2502	
% of Correct Classifications		12		
Sample size		434		

^a Dependent variable = 1 (0) if performance measure was (was not) adjusted (A).

^b Div_{i} are the division indicators.

Probit Model with Number of Managers as an Additional Control Variable^a

	Pred	<u>Model</u>		
	<u>Sign</u>	Coeff.	<u>z-stat</u>	
Intercept		-5.697	-1.79*	
SensitivityCongruity	-	-0.036	-0.18	
Noise	+	0.389	2.15**	
CompletenessVerifiability	-	-0.341	-2.18**	
Manipulability	+	-0.081	-0.49	
InfluenceActivity	+	1.820	3.37***	
Favoritism	+	0.899	2.81***	
Planning		-0.684	-2.92***	
Reporting		-0.251	-1.24	
ProfessionalDevelopment		-0.501	-1.66*	
Div ₂ ^b		2.497	3.40***	
Div ₄		-3.116	-2.84***	
Div5		0.066	-0.14	
Div9		-3.228	-2.71***	
Div ₁₀		0.400	0.54	
Div ₁₁		-2.691	-3.27***	
Div ₁₂		0.389	-0.87	
DivisionPerformance		-0.106	-0.76	
Level		0.132	1.13	
NumberManagers		-0.357	-1.81*	
$\mathbf{P}_{\mathrm{result}} \mathbf{P}^2$			116	
		0.1	110	
% of Correct Classifications		72	.58%	
Sample size		434		

^a Dependent variable = 1 (0) if performance measure was (was not) adjusted (A).

^b Div_i are the division indicators.

BIBLIOGRAPHY

- Anderson, S.W., H.C. Dekker, and K.L. Sedatole. 2007. An empirical examination of negotiated goals and performance-to-goal following the introduction of an incentive bonus plan. Working Paper.
- Antle, R. and Demski, J.S. 1988. The controllability principle in responsibility accounting. *The Accounting Review* 63 (4): 700-18.
- Argyris, C. 1964. Integrating the individual and the organization. Wiley: New York.
- Arvey, R.D. and K.R. Murphy. 1998. Performance evaluations in work settings. Annual Review of Psychology 49:141-68.
- Arya, A., J. Glover, B. Mittendorf, and L. Ye. 2005. On the use of customized versus standardized performance measures. *Journal of Management Accounting Research*: 7-21.
- Bailey, W. J., G. Hecht, and K. L. Towry. 2006. Dividing the pie: Do managers fully incorporate non-contracted information into full and partial discretionary bonus allocations? Working Paper.
- Baiman, S. and M. V. Rajan. 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review* 70(4): 557-79.
- Baker, G.P. 1990. Pay-for-performance for middle managers: causes and consequences. *Journal of Applied Corporate Finance* 3(3): 50-61.
- Baker, G. P., R. Gibbons, and K. J. Murphy. 1994. Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics* 109(4): 1125-56.
- Baker, G. 2000. The use of performance measures in incentive contracts. *The American Economic Review* 90(2): 415-20.
- Banker, R. D. & Datar, S. M. 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* 27: 21-39.
- Bol, J.C. 2005. Subjective performance evaluation. Working paper.
- Bretz, R.D., G.T. Milkovich, and W. Read. 1992. The current state of performanceappraisal research and practice – concerns, directions, and implications. *Journal* of Management 18(2): 321-52.
- Bushman, R. M., R. J. Indjejikian and A. Smith. 1996. CEO compensation: The role of individual performance evaluation, *Journal of Accounting and Economics*, 21(2): 161-93.
- Coronado, F. 2007. Choosing performance measures for incentive compensation. Unpublished Manuscript.
- Coronado, F. and R. Krishnan. 2007. Task complexity and incentive contracting an empirical investigation. Working paper.
- Datar, S., S. Kulp, and R. Lambert. 2001. Balancing performance measures. Journal of Accounting Research: 75-92.

- Evans, J. H., R. L. Hannan, R. Krishnan, and D. V. Moser. 2001. Honesty in managerial reporting. *The Accounting Review*, 76: 537-560.
- Feltham, J. & Xie, J. (1994). Performance measure congruity and diversity in multitask principal/agent relations. *The Accounting Review*: 429-453.
- Ferris, G. R. and T. A. Judge. 1991. Personnel/human resources management: A political influence perspective, *Journal of Management* 17(2): 447-88.
- Fisher, J.G. Frederickson, J.R., and S.A. Peffer. 2000. The budgeting process: an experimental investigation of the effects of negotiation. *The Accounting Review*, January 75: 93–114.
- Fisher, J.G. Frederickson, J.R., and S.A. Peffer. 2002. The effect of information asymmetry on negotiated budgets: an empirical investigation. *Accounting Organizations and Society*, 27: 27–43.
- Fisher, J. G., L. A. Maines, S. A. Peffer, and G. B. Sprinkle. 2002. Using budgets for performance evaluation: effects of resource allocation and horizontal information asymmetry on budget proposals, budget slack, and performance. *The Accounting Review* 77(4): 847-865.
- Fisher, J.G., L.A. Maines, S.A. Peffer, and G.B. Sprinkle. 2005. An experimental investigation of employer discretion in employee performance evaluation and compensation. *The Accounting Review* 80(2): 563-83.
- Frederickson, J.R. 1992. Relative performance information: The effects of common uncertainty and contract type on agent effort. *The Accounting Review*: 67(4) 647-69.
- Gibbs, M., K. A. Merchant, W. A. Van der Stede, and M. E. Vargus. 2004. Determinants and effects of subjectivity in incentives, *The Accounting Review* 79(2): 409-36.
- Govindarajan, V. 1984. Appropriateness of accounting data in performance evaluation: an empirical examination of environmental uncertainty as an intervening variable. *Accounting, Organizations, and Society* 9: 125-36.
- Haag, S., Cummings, M., McCubbrey, D., Pinsonneault, A., & Donovan, R. 2006. Management Information Systems For the Information Age (3rd Ed.). Canada: McGraw Hill Ryerson.
- Hayes, R. M. and S. Schaefer. 2000. Implicit contracts and the explanatory power of top executive compensation for future performance. *Rand Journal of Economics* 31(2): 273-93.
- Holmstrom, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 10(1): 74-91.
- Holmstrom, B. and P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design, *Journal of Law, Economics and* Organization 7(2): 24-53.

- Hoppe, F. and F. Moers. 2007. Discretion and the complexity of simple bonus contracts. Working Paper.
- Hopwood, A.G. 1972. An empirical study of the role of accounting data in performance evaluation. *Journal of Accounting Research* 10 (Supplement): 156-82.
- Ilgen, D.R., and J.L. Favero. 1985. Limits in generalization from psychological research to performance appraisal processes. Academy of Management Review 10(2): 311-21.
- Ilgen, D.R. and J.R. Hollenbeck. 1991. The structure of work: job design and roles. In *Handbook of Industrial and Organizational Psychology*, Vol. 2, edited by Dunnette and Hough. Palo Alto, CA.
- Ittner, C.D., and D.F. Larcker. 1998. Innovations in performance measurement: trends and research implications. *Journal of Management Accounting Research* 10: 205-38.
- Ittner, C. D., D. F. Larcker, and M. Meyer. 2003. Subjectivity and the weighting of performance measures: evidence from a balanced scorecard. *The Accounting Review* 78 (3): 725-58.
- Kachelmeier, S. 1996. Do cosmetic reporting variations affect market behavior? A laboratory study of the accounting emphasis on unavoidable costs. *Review of Accounting Studies* 1: 115-140.
- Kaplan, R. S. and D. P. Norton. 1996. *The Balanced Scorecard*. Boston, MA: Harvard Business School Press.
- Kaplan, R. S. and D. P. Norton. 2001. *The Strategy-Focused Organization*. Boston, MA: Harvard Business School Press.
- Krishnan, R., J. Luft, & M. D. Shields. 2005. Effects of accounting-method choices on subjective performance measure weighting: Experimental evidence on precision and error covariance. *The Accounting Review* 80(4): 1163-92.
- Kwon, Illoong. 2005. Endogenous favoritism in organizations. Working paper, University of Michigan.
- Lipe, M. G. and S. E. Salterio. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review* 75(3): 283-98.
- Longenecker, C. O., H. P. Sims, and D. A. Gioia. 1987. Behind the mask: The politic of employee appraisal. *The Academy of Management Executive* 1(3): 183-93.
- MacLeod, B.W. 2003. Optimal contracting with subjective evaluation. The American Economic Review 93(1): 216-40.
- MacLeod, B. W. and D. Parent. 1999. Job Characteristics and the form of compensation. *Research in Labor Economics* 18: 177-242.

- Matsumura, E.M., and J.Y. Shin. 2006. An empirical analysis of an incentive plan with relative performance measures: Evidence from a postal service. *The Accounting Review*: 81(3): 533-66.
- McDonald, P. 2003. Staffing today's internal audit function: audit executives need a realistic strategy for obtaining top talent to handle growing demands. *Internal Auditor* 60(5): 46-52.
- Merchant, K.A. 1989. *Rewarding Results*. Harvard Business School Press: Boston, Massachusetts.
- Merchant, K.A., W.A. Van der Stede, and L. Zheng. 2003. Disciplinary constraints on the advancement of knowledge: the case of organizational incentive systems. *Accounting, Organizations, and Society* 28: 251-86
- Milgrom, P.R. 1988. Employment contracts, influence activities, and efficient organization design. *Journal of Political Economy* 96(1): 42-60.
- Milgrom, P., and J. Roberts. 1988. An economic approach to influence activities in organizations. *American Journal of Sociology* 94:154-79.
- Milgrom, P. and J. Roberts. 1992. *Economics, organization and management*. Prentice-Hall: New Jersey.
- Milkovich, G.T. and A.K. Wigdor. 1991. Pay for Performance: Evaluating Performance Appraisal and Merit Pay. National Academy Press, Washington, D.C.
- Moers, F. 2005. Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations and Society* 30(1): 67-80.
- Moers, F. 2006. Performance measure properties and delegation. *The Accounting Review*: 81(4): 897-924.
- Murphy, K. J. and P. Oyer. 2003. Discretion in executive incentive contracts: Theory and evidence. *Working Paper*.
- Murphy, K.R. and J.N. Cleveland. 1991. Performance Appraisal: An Organizational Perspective. Boston: Allyn and Bacon.
- Murphy, K.R. and J.N. Cleveland. 1995. Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives. Thousand Oaks: Sage Publications.
- Orey, M. 2006. White Men Can't Help It. Business Week, May 15, 2006: 54-7.
- Prawitt, D.F. 2003. Managing the internal audit function. In A.D. Bailey, A.A. Gramling, and S. Ramamoorti (Eds.) *Research Opportunities in Internal Auditing* (171-229). The Institute of Internal Auditors Research Foundation.
- Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37(1): 7-63.
- Prendergast, C. 2002. The tenuous trade-off between risk and incentives. *Journal of Political Economy* 110(5): 1071-1102.

- Prendergast, C. and R. H. Topel. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37(2-3): 355-65.
- Prendergast, C. and R. H. Topel. 1996. Favoritism in organizations. Journal of Political Economy 104(5): 958-78.
- Rajan, M.V. and S. Reichelstein. 2006. Subjective performance indicators and discretionary bonus pools. *Journal of Accounting Research* 44(3): 585-618.
- Taylor, E., and R. Wherry. 1951. A study of leniency of two rating systems. *Personnel Psychology* 4:39-47.
- Tziner, A., G.P. Latham, B.S. Price, and R. Haccoun. 1996. Development and validation of a questionnaire for measuring perceived political considerations in performance appraisal. *Journal of Organizational Behavior* 17: 179-90.
- Van der Stede, W.A., C.W. Chow, and T.W. Lin. 2006. Strategy, choice of performance measures, and performance. *Behavioral Research in Accounting* 18: 185-205.
- Waller, W.S. 1995. Decision-making research in management accounting: Return to behavioral-economics foundations. In *Judgment and Decision-Making Research in Accounting and Auditing*, edited by R.H. Ashton and A.H. Ashton. New York, NY: Cambridge University Press.
- Wayne, S.J., R.C. Liden, I.K. Graf, and G.F. Ferris. 1997. The role of upward influence tactics in human resource decisions. *Personnel Psychology* 50: 979-1006.

