Firms' Commitment to Mitigate the Ratchet Effect under Annual Bonus Plans with Multiple Performance Measures

BYUNG HYUN CHOI
LG Economic Research Institute

JONGHWAN KIM*
Università Bocconi

KENNETH A. MERCHANT
University of Southern California

April, 2014

Preliminary and incomplete.

PLEASE DO NOT QUOTE OR CITE WITHOUT PERMISSION.

* Corresponding author: Department of Accounting, Università Bocconi, Milan 20136, Italy. E-mail: Jonghwan.Kim@unibocconi.it.

We gratefully acknowledge the support of the company that provided our research data which must, unfortunately, go unnamed.

Firms' Commitment to Mitigate the Ratchet Effect under Annual Bonus Plans with Multiple Performance Measures

ABSTRACT

Prior literature has documented ratcheting targets and a potential adverse incentive problem it may cause, and, as a remedy to the problem, proposed a firm's commitment to incomplete use of past performance in setting future targets. In this paper, we examine whether a firm makes such a commitment and whether it behaves differently depending on the past target achievement level, the type of a performance measure, and the importance weight. Using a confidential dataset that consists of 1,208 performance-measure-year observations with the complete set of information including targets, actual performance, and importance weights for all performance measures used in the short-term incentive plan, we find that (1) the firm makes a credible commitment; (2) the firm's target revision behavior, however, varies with the previous target achievement level—rewarding past good performance, eliminating excess good luck, and compensating for excess bad luck; (3) the firm's behavior also varies with the type of performance measure—making stronger commitments for good financial and common measures while penalizing bad performance in the measures; and (4) contrary to our prediction, importance weights do not affect the firm's target revision behavior. Overall, generally consistent with prior literature, our study adds empirical evidence with novel findings, making the best use of the unique dataset.

Keywords: *performance targets, performance measures, inter-temporal performance manipulation, target difficulty, adverse incentive, the ratchet effect*

Data Availability: *Data used in this study are derived from proprietary sources.*

Firms' Commitment to Mitigate the Ratchet Effect under Annual Bonus Plans with Multiple Performance Measures

I. INTRODUCTION

Many firms reward their managers based on the performance relative to targets that have been set prior to or at the beginning of a performance period (Murphy, 2001). The use of performance targets allows firms to provide workers with high-powered incentives while providing a competitive, expected compensation (Milgrom and Roberts, 1992; Murphy, 2001). Despite the benefit, firms that use past performance to identify workers' productivity and to set future targets often face a serious adverse incentive problem—so-called the "ratchet effect" (Milgrom and Roberts, 1992; Weitzman, 1980). Contracting theory recommends that, to mitigate the problem, firms should not reflect the past performance fully on future targets (Laffont and Tirole, 1993; Milgrom and Roberts, 1992). Only a few studies, however, have attempted to test whether firms make the commitment not to use all available information in target setting (Bouwens and Kroos, 2011; Indjejikian et al., 2012; Indjejikian and Nanda, 2002).

Our objective of this study is to provide empirical evidence about a firm's target-setting behavior in a multi-measure performance evaluation system. In particular, this paper examines how a firm's commitment to incomplete use of past performance varies with the level of past target achievement (i.e., performance relative to a target) and with performance measure characteristics. To address our research questions, we analyze a confidential internal dataset containing the annual performance evaluation of business unit managers in a conglomerate. The dataset consists of 1,208 performance-measure-year observations for 133 business unit-years from 2005 to 2009.

In the firm of our study, we find evidence supporting that (1) the firm demonstrates such a commitment only to the managers whose performance was above a target, (2) tries to eliminate a luck factor that contributes to unexpected performance far exceeding a target by a large margin, and (3) resets a target difficulty level such that poor performance stops recurring for consecutive years. Additionally, our investigation of the effects of performance measure types and importance weights on a firm's target revision behavior shows that performance relative to a target in financial or common measures are more likely to persist than nonfinancial or unique measures, and that importance weights do not affect a firm's target revision practice.

With the novel findings from the unique dataset, we contribute to the literature in several ways. First, the access to a proprietary dataset allows us a direct and comprehensive look at a firm's target revision practice. Our dataset contains a complete set of data used for annual bonus determination, including the full set of performance measures that are included in the firm's short-term performance plan, their targets and actual performance outcomes, and importance weights. Without such data, prior studies in this area instead depend on inferred performance (e.g., bonus payout in Indjejikian and Nanda, 2002), performance solely in one type of performance measure (e.g., Bouwens and Kroos, 2011; Leone and Rock, 2002), or survey (e.g., Indjejikian et al., 2012). Unlike prior studies, our study uses performance *per se* in different types of measures.

Second, to our knowledge, this is the first empirical study that compares target-setting behaviors for a variety of performance measures. Research on performance targets has been centered on budgets, i.e., accounting or financial performance targets. Little research investigates targets in other performance measures. However, knowledge about

target-setting only in the budgeting contexts is incomplete to explain the behavior of firms with multiple performance measures in their incentive system. Our study fills the void, examining the target-setting behavior of a firm with a multi-measure performance evaluation system. Specifically, it classifies performance measures into financial vs. nonfinancial and common vs. unique measures and provides first evidence that a firm shows different behaviors in revising targets of different performance measures.

Third, we also report managers' target-achieving behavior in a multiple performance measure setting. Specifically, we document a discontinuity at a benchmark in different types of performance measures other than accounting earnings. Despite the well-documented "discontinuity at zero earnings" in earnings management literature (e.g., Burgstahler and Chuk, 2012; Burgstahler and Dichev, 1997; Burgstahler and Eames, 2006; Degeorge et al., 1999; Hayn, 1995) and our general expectation about its manifestation in a variety of performance benchmarks, it has rarely been reported for performance measures other than accounting earnings. This paper first reports the phenomenon even in other types of performance measures and thereby contributes to the literature.

In the next section, we review related prior literature and develop research hypotheses. In Section 3, we describe our research setting, focusing on notable features of the research site's target-setting process and short-term incentive plans. Section 4 describes our sample and measures. In Section 5, we present our empirical results including descriptive evidence and tests of hypotheses. Finally, Section 6 summarizes our findings, discusses limitations of our study, and offers future research potentials.

II. THEORY AND HYPOTHESIS DEVELOPMENT

Target setting as part of planning practice is a critical element of firms' management control systems. As an important result control device, targets are set to motivate workers. To set a target that produces motivational effects as intended, targets should be set such that they are "neither too easy nor too difficult to achieve" (Merchant and Van Der Stede, 2007; p. 330). To figure out targets with the level of difficulty, firms need to identify worker's productivity first. To that end, firms study task-specific efficiency, compare a worker's performance with that of others in similar jobs, or use the past performance of the same worker in the same job (Milgrom and Roberts, 1992).

When past performance is used as a performance benchmark, future targets are often ratcheted up for good past performance at least under budgeting contexts (Choi and Lee, 2004; Indjejikian et al., 2012; Leone and Rock, 2002; Weitzman, 1980). This creates an incentive problem: the "ratchet effect" (e.g., Bouwens and Kroos, 2011; Milgrom and Roberts, 1992; Murphy, 2001; Weitzman, 1980). The ratchet effect refers to an undesirable outcome of target ratcheting;[1] specifically, a situation where workers opt to stop short of exerting their full effort once current targets are achieved (Milgrom and Roberts, 1992; Weitzman, 1980). Under target-ratcheting circumstances, managers with performance close to a target level have conflicting incentives—facing a tradeoff between current rewards from better current performance and future losses from the assignment of ratcheted (i.e., more challenging) targets. On one hand, managers need to meet the target to earn bonuses. One the other hand, once they meet the target, they stop short of exceeding targets. Anticipating target ratcheting and perceiving it as an undue

[1] Milgrom and Roberts (1992; p. 233) defined it as "the tendency for performance standards to increase after a period of good performance." This definition corresponds to "the ratchet principle" in Weitzman (1980) and Leone and Rock (2002).

penalty for good performance, managers try to shift the timing of performance realization to the future or reduce their effort level so that they can maximize their total compensation over the multiple-period term through inter-temporal performance management (Bouwens and Kroos, 2011; Holthausen et al., 1995; Indjejikian et al., 2012; Indjejikian and Nanda, 2002; Leone and Rock, 2002; Milgrom and Roberts, 1992; Murphy, 2001; Weitzman, 1980).

The evidence of such behavior is documented in Leone and Rock (2002), Murphy (2001), and Bouwens and Kroos (2011). Leone and Rock (2002) report that managers make income-decreasing accounting choices as to discretionary accruals when a positive earnings deviation from a target is considered transitory. More direct evidence of performance management in a planning cycle is documented in Murphy (2001) and Bouwens and Kroos (2011). With quarterly performance data, they find that managers with favorable year-to-date performance reduce the level of their effort exertion in the last quarter of a fiscal year.

Firms' Commitment: Incomplete Use of Available Information

Despite the adverse impact of target ratcheting, setting targets based on past performance is still prevalent. Further, many firms prepare budgets and pay bonuses computed based on performance relative to the budgets (e.g., Libby and Lindsay, 2010; Murphy, 2001; Umapathy, 1987). Then, the next question arises regarding a firm's response to managers' behavior to avoid beating targets; "do firms ignore or fix the problem?" Theory suggests firms' commitment not to fully incorporate information about a manager's productivity revealed in the past period performance into the next period's

target (e.g., Indjejikian and Nanda, 2003; Laffont and Tirole, 1993; Milgrom and Roberts, 1992).[2]

Consistent with what theory suggests, empirical findings from a few, recent studies show that firms ratchet targets but do not use full information available in the past performance (Bouwens and Kroos, 2011; Indjejikian et al., 2012; Indjejikian and Nanda, 2002). To show firms' commitment to incomplete use of past performance in setting future targets, the literature uses the serial correlation of target achievements between consecutive years. Specifically, Indjejikian and Nanda (2002) argue that if firms incorporate information in the past performance fully into next period's targets, target achievement level will not be serially correlated. To test their hypothesis, Indjejikian and Nanda (2002) compare actual and target bonuses, instead of performance, of a large number of executives, and find a positive serial-correlation of annual bonuses; an executive who earned a bonus in a previous period is more likely to earn a bonus in the following period. Bouwens and Kroos (2011) report similar, but differently motivated, results using store managers' quarterly sales performance in a Dutch retailer. They find that managers who reduce sales in the final quarter of a fiscal year (i.e., with their annual sales target met but not by a large margin) are more likely to achieve the following year's targets than the others. The finding suggests that managers are successful in creating slack in a target even before the next year's planning process. More importantly, it also suggests a firm's commitment to hold the line even after good performance because managers' slack creation can be sustained under the tacit consent of a firm.[3]

[2] Other remedies to the problem include, but not limited to, job rotation (Milgrom and Roberts, 1992), the use of aggregate performance measures (Indjejikian and Nanda, 1999), and the use of external performance standards (Murphy, 2001).

[3] We preclude the possibility of ineffective management control—i.e., a firm's inability to identify slack.

In sum, the findings in prior empirical studies consistently suggest firms' commitment to incomplete use of past performance in target setting. Our first hypothesis predicts a positive correlation of the achievability of a target between two consecutive fiscal years.

Hypothesis 1. *The likelihood (extent) of target achievement is positively associated with (the extent of) target achievement in the previous year.*

In addition to a significant serial correlation of abnormal bonus earnings between consecutive years, Indjejikian and Nanda (2002) document firms' asymmetric target revision. They find that the serial correlation is insignificant when previous targets are missed. In other words, firms' incomplete reflection of past performance on future targets is limited only when previous targets are achieved. In a similar vein, Indjejikian et al. (2012) report that firms revise the target difficulty more favorably (unfavorably) for well-performing (poorly-performing) managers.

These findings about firms' asymmetric target revision are also consistent with firms' target difficulty setting practices—i.e., incentives to allow more slack or to require higher targets—discussed in Merchant and Manzoni (1989). Specifically, firms tend to be generous to well-performing managers' slack creation to retain efficient managers by assuring them with certain economic rents. Merchant and Manzoni (1989) describe managers' incentives to lower target difficulty: for example, to avoid the loss of goal commitment of those well-performing managers, to make less control interventions, to allow discretion, or to ensure a competitive compensation package. On the other hand, firms may require more difficult targets to communicate their dissatisfaction with managers' performance (Merchant and Manzoni, 1989).

Firms' target revision favoring only well-performing managers, or conversely disfavoring poorly-performing managers, may occur in two forms: fair incorporation of productivity information vs. penalization. First, targets may be adjusted on a fair and neutral basis. In this case, a firm utilizes information about a manager's productivity disclosed in the previous year's performance to set a new target. Target revision of this type creates no serial correlation of performance to a target between years. Second, poor performance may be penalized such that a new target is set as difficult as the previous target. Hence, it allows a positive association of performance to a target between years. To accommodate firms' different attitude toward managers with different performance, we hypothesize it in two competing alternatives about a moderating effect of performance.

Hypothesis 2a. *When a target is not achieved in the previous period, the likelihood (extent) of target achievement is not related to target achievement in the previous period.*

Hypothesis 2b. *When a target is not achieved in the previous period, the likelihood (extent) of target achievement is positively related to target achievement in the previous period.*

Realized performance that far exceeds a target may indicate either (1) the presence of uncontrollable "luck"—i.e., unexpected positive exogenous news or (2) the absence of managers' effort reduction, or both.[4] If the luck is completely transitory, the same level of target achievement in the following period can hardly be sustained. By definition, this indicates no serial correlation of target achievements of consecutive years. If such a positive business environment is expected to continue in the future, a firm may adjust the luck factor by requiring more challenging (or less achievable) targets

[4] Given the managers' "meet but not beat a target" incentive (Anderson et al., 2010; Bouwens and Kroos, 2011; Leone and Rock, 2002), the former is more likely.

(Merchant and Manzoni, 1989). This also suggests an insignificant correlation of target achievements between consecutive years. This will result in non-linearity in the relationship at a certain level of performance that is far above a target; excess achievements may not be as rewarding as just-above-target achievements, or may be even penalized. We hypothesize the relationship in a null form.

Hypothesis 3. *When a target is exceeded by a large margin, the likelihood (extent) of target achievement is not related to target achievement in the previous period.*

Considering managers' incentive to achieve targets and ability to build budgetary slack, realized performance far below a target[5] is likely to be an outcome of unexpected, exogenous negative shocks. As huge negative shocks are often uncontrollable, firms may take them into consideration in setting next period's targets. A firm's adjustment for uncontrollable huge negative impact would reset a target achievability level to the firm's conventional level of target achievability. This will result in no serial correlation of performance to a target between years.

Hypothesis 4. *When a target is missed by a large margin in the previous period, the likelihood (extent) of target achievement is not related to target achievement in the previous period.*

Effects of Performance Measures Used in Bonus Formula

Many firms include alternative performance measures as well as financial/accounting-based targets in their incentive plans (e.g., Ittner et al., 1997; Kaplan and Norton, 1992). Despite the wide use of non-budget targets in practice, prior literature has been silent about how these targets are set and revised.

[5] By "performance far below a target," we refer to a huge divergence from a target level that is out of the conventional range. The definition of conventional range is subjective and discretionary. In this paper, we define performance far below a target as performance to a target less than 80%. Table 2 reports that 12.4% of the full sample are in this category.

Different typologies may sort performance measures in many different ways. Prior literature studying multiple performance measures (e.g., Balanced Scorecards) has often used common vs. unique (e.g., Lipe and Salterio, 2000) and financial vs. non-financial (e.g., Ittner et al., 1997) classifications. For example, earnings is a common accounting-based measure, while the rate of on-time delivery is a less common,[6] non-financial performance measure.

Use of alternative performance measures involves different degrees of information asymmetry and different importance weights. Compared to unique or non-financial measures, communicating, understanding and reviewing the level of targets and their achievability of common or financial/accounting-based measures require relatively less cognitive effort, which results in firms' or superiors' bias toward common measures as a communication means for evaluation purposes (Libby et al., 2004; Lipe and Salterio, 2000; Slovic and MacPhillamy, 1974).

The superior's bias—placing more weight on common measures—may increase or decrease the serial correlation of target achievements between years. On one hand, superiors' greater emphasis on common measures may worsen the potential adverse impact of target ratcheting. The preference manifests itself in greater importance weights on common or financial measures in a bonus formula (Lipe and Salterio, 2000). Accordingly, greater influence of a performance measure on a bonus prompts managers' stronger incentive to achieve the target. Managers, thus, may engage in inter-temporal

[6] Not all firms or business units have these types of non-financial, disaggregate performance indicators as key performance measures for bonuses. In the balanced scorecard context, business units have some unique measures that are specifically tailored to a business unit's strategy and operation as well as some common, generic measures (e.g., Kaplan and Norton, 1996; Lipe and Salterio, 2000).

performance management or slack creation whenever it is possible.[7] As such, a measure with a large importance weight may be more prone to the ratchet effect than that with a small weight. This is where firms' commitment is elicited.

On the other hand, superiors' preference of common or financial measures may have the opposite effect. Specifically, superiors' preference of common or financial measures as a communication means makes it more difficult for managers or subordinates to hide their private, local information communicated with such measures. In other words, the measures provide superiors with greater ability to detect slack, which as a result can frustrate managers from creating slack (Lal et al., 1996; Merchant, 1985). With less information asymmetry allowed in measures, managers can be less involved in creating slack while superiors may depend less on the past performance to set future targets. If this is the case, firms may not be concerned about the ratchet effect, no commitment is needed, and thus no serial correlation would be observed.

The discussion so far produces inconclusive predictions about the effects of the choice of performance measure on a firm's commitment to deemphasize past performance in a target setting process. So, we develop hypotheses in a null form regarding the effects of performance measure choice.

Hypothesis 5. *The relationship of target achievements between two consecutive years does not vary with the type of performance measures.*

As performance measures carry different importance weights in bonus determination formula, the effects of importance weights may well be examined together with the effects of performance measure type. Unlike the effects of the type of

[7] Low information asymmetry does not indicate that all slack is detectable and, even when it is detected, not all the detected slack is removed.

performance measures, the previous discussion as to the effects of importance weights renders an unequivocal prediction.

Hypothesis 6. *The relationship of target achievements between two consecutive years is stronger for measures with a large importance weight than for those with a small importance weight.*

III. RESEARCH SETTING

The Research Site

Our research site is a large conglomerate based in South Korea, which we will refer to as KC (short for Korean Conglomerate). KC is a huge conglomerate governing approximately 40 companies, as of 2010, in a variety of industries including electronics, chemicals, and telecommunications. KC's overall annual sales in 2010 is well over $US 100 billion and its employees are over 200,000 world-wide. KC's management control system has been developed to facilitate decentralized control. It is characterized as follows. First, despite the high latitude of autonomy granted to business units, they are arranged in a strict hierarchy. Second, budget and management-by-objectives (MBO) constitute important components of KC's management control system, and affect managers' incentives and behaviors to a great extent. Finally, the primary goal of KC's performance evaluation system is to secure fairness and objectivity of evaluations and to provide its managers with differential compensations corresponding to their performance.

The Budgeting, Target-setting, and Performance Evaluation Process

The Timeline

Figure 1 illustrates KC's planning (i.e., budgeting and target-setting) and performance evaluation process. In September, an annual planning process begins with business entity's preparation of budgets and performance targets. After the initial

submission, the targets are evaluated and modified through communications back and forth along the hierarchy. KC's in-house consulting organization is actively involved in budget evaluations, supporting the KC's top management. KC finalizes the budgets and performance targets of large business entities by the end of October.

Insert Figure 1 about here.

Budgeting, Target-Setting Reviews, and Managers' Behaviors

Budgets and performance targets are thoroughly reviewed by the conglomerate's in-house consultants. These experts' primary job in the budget/target review is to assist the conglomerate-level management's decisions by collecting and providing likely private information about businesses and industries and, as a result, to reduce information asymmetry lying between managers of different organization levels (i.e., KC top management and business entity managers). They are expected to have as good, if not complete, local knowledge as business entity managers. The presence of these expert gatekeepers provides backgrounds for managers' behavior in the budgeting and target-setting process.

First, facing experts who support superiors' review, business entity managers are less able to build huge slack in their budgets and targets as compared otherwise. Accordingly, managers are pressured to bring up "reasonable" targets to persuade these experts and ultimately their supervising managers at a higher-organization level. Second, to develop reasonable targets, managers often base their estimates of future performance on the past performance, because they well understand that targets below the past performance are hardly approved. This is consistent with a way to set a performance standard in a reasonably objective manner provided by Milgrom and Roberts (1992; p.

233). In sum, the described budgeting process provides a fair background for target ratcheting.

In addition to managers' concern about the presence of budget review experts, managers have other good reasons to make deliberate decisions about their budgets and performance targets. First, once finalized, the budgets and targets are rarely renegotiated with exceptions of extraordinary situations such as a natural disaster and a severe recession shock. Second, they will make standards for the next year's performance evaluation of business entity managers for bonus determination. Considering the inflexibility and economic substance, managers set targets very carefully and commit efforts to negotiating favorable budgets.

The Incentive System

KC runs two separate incentive programs; short-term and long-term incentive plans. The short-term incentive plan (STIP) has been in place far ahead of the beginning of the earliest sample period (i.e., 1999) while a cash-based long-term performance plan (LTPP) was introduced in 2006 for the first time. Stock options were granted to a limited number of high-rank executives and abolished before 2003.[8]

This study focuses on KC's STIP. With STIP, KC rewards its managers based on two equally weighted categories of performance measures: summary financial measures and key performance indicators (KPIs). Typical summary financial measures that are used by all business units are sales and operating profits. On the other hand, KPIs are selected to be informative of the performance of core strategic tasks that are unique to

[8] The stock option plan period overlaps the sample period of one of our datasets. However, the options were granted to only a few highest-rank executives. Therefore, the business unit managers in the dataset whose hierarchical ranks do not qualify option grants are hardly affected by the then-existing stock option plan.

each business unit. KPIs encompass a variety of performance indicators including specific financial indicators such as cost reductions and sales of key products, and non-financial measures such as market share and brand recognition indexes. Importance weightings of each constituent performance measure are negotiated as part of the annual planning process.

Insert Figure 2 about here.

Target achievement, or performance relative to target (i.e., actual/target), in each defined performance measure is converted into the evaluation ratings of Overachieved, Achieved, or Missed. Relative performance to peers or previous periods is classified into Outstanding, Par, or Below Standard. Considering target achievement and relative performance together, the comprehensive evaluation on each performance measure is turned into a nine-point scale score. Then, scores are weight averaged within each category (i.e., summary financial measures and KPIs). The evaluation score in each performance dimension produces a five-point scale grades (S,[9] A, B, C, or D). The grades in both dimensions are used to define the bonus amounts to be paid to each manager. For example, executives who receive S grades in both dimensions earn an annual bonus of ten times their *monthly* salary (i.e., 500% from each criterion) while those with D grades in both dimensions earn no short-term bonus.

IV. RESEARCH DESIGN

Data

[9] Stands for "Superior."

We conduct our analysis using a set of data about business entity managers' annual performance evaluations. Originally, the dataset consists of 1,214 performance measure-years that the conglomerate used for annual performance evaluation of its twenty-nine business entities (and their managers) from 2005 to 2009. Out of these observations, we exclude six observations; five observations for which proper importance weights are not provided or imputed, and one observation of which the characteristic (e.g., financial vs. non-financial and common vs. unique) is unidentifiable. As a result, the final dataset include 1,208 observations, each of which is provided with all necessary, relevant information for this study such as a performance target, actual performance, importance weight.

Variables

Performance Relative to Target (PRT). Performance-relative-to-target (*PRT*) is the extent to which a specific target is achieved by a manager, capturing an *ex post* measure for target difficulty at the same time. *PRT* is computed as the ratio of a realized performance relative to a target denoted in a contracted measure, written in the following formula:

$$PRT_{it} = \frac{realized\ performance_{it}}{target_{it}},$$

where *i* represents a performance measure used in STIP, *t* denotes a year. A *PRT* that is greater than (equal to, less than) one indicates that a target is achieved (exactly met, not achieved). As mentioned above, the measure captures the relative difficulty level of a performance target. For example, a target with *PRT* of 120% indicates not only that realized performance exceeds the target by 20% but also that meeting the target is more achievable than another performance target with *PRT* of 105% or 95%.

Target Achievement. For performance targets of which *PRT* is greater than or equal to one, an indicator variable $Achieved_t$ is assigned one and zero otherwise. In addition, we define *overachievement* (*underachievement*) as performance exceeding (falling below) 120% (80%) of a target level, and assign one to an indicator variable, $PRT \geq 120_t$ ($PRT < 80_t$) accordingly.

Performance Measure Types. Performance measures are categorized in two ways: financial vs. nonfinancial and common vs. unique. The first categorization is simple and straightforward. Financial measures include, for example, sales (growth), profit (margin), and EBIT(D)A. All the other measures that are hardly considered financial are, by definition, classified into a nonfinancial category. To separate common and unique measures, we first consider business entity or task specific performance measures—for example, production capacity for a certain line of product and the degree of supplier diversification—that are rarely used in other business entities. Using this method, 243 out of 1,208 measure-years are classified as unique. In addition to these unique measures, we identify performance measures of relatively less frequently used types than others and add them to the unique measure category.[10] This identifies additional 156 measure-years to make the total of 399 unique-measure-years.

Importance Weight. Importance weights on performance measures contributing to bonus calculation are provided as part of dataset. As explained in a previous section,

[10] Table 1 explains how we identify additional unique measures. Specifically, we group eleven classes of performance measures by similarity: for example, measures of profit, sales, customer satisfaction, etc. Then, we compute the median frequency of adoption without the 238 previously identified, company-specific, unique performance measures. It is 58 times for which market share type measures are used. With the median frequency of adoption, we divide measures into more frequently adopted (i.e., common) and less frequently adopted (i.e., unique) types of measures. The less frequently adopted types of measures are human resource management, research and development, brand recognition, cost management, and initiatives. The more frequently adopted measures include profit, sales, (business) composition, KC's employee survey results, customer satisfaction, and market share.

performance for an STIP bonus is evaluated based on two dimensions—summary financial measures and KPIs. Each dimension contains a set of performance measures whose importance weights sum up to 50%.[11]

V. EMPIRICAL RESULTS

Descriptive Statistics

Use of Performance Measures in STIP. Table 2 presents descriptive statistics of our sample. Panel A of Table 2 shows that the observations are reasonably balanced across the sample period from 2005 to 2009. On average, a business entity uses nine performance measures in their short-term incentive plans. Panel B provides the composition of the performance measure sample. Out of the 1,208 performance measure-years, 237 measure-years (19.6%) are isolated observations; they are used (1) only for one year or (2) with a time gap. Therefore, the analysis of the inter-temporal target revision uses the other 971 observations whose *PRT* data are available for at least two consecutive years, excluding the isolated observations. Speaking of the type of measures, there are 601 (607) financial (nonfinancial) performance measure-years accounting for 49.8% (50.2%) of the sample, while there are 809 common measure-years (67.0%) and 399 unique measure-years (33.0%). The panel also reports that the mean (median) importance weight is 10.8% (10%).

Insert Table 2 about here.

Distribution of Actual Performance to Target. As for the performance, Panel B shows that performance targets are more likely to be achieved than not; they are achieved

[11] See Figure 2.

with 59.1% of probability. The mean of *PRT* is 108.6% while its median (101.1%) is slightly above 100%. Further, the probabilities of being overachieved (i.e., *PRT*≥120%) and underachieved (i.e., *PRT*<80%) are 15.7% and 12.4% respectively. Conversely, about 72% of realized performance outcomes lie between 80% and 120% of target levels.

Insert Figure 3 about here.

Figure 3 confirms the dense distribution of actual performance immediate around target level performance. More importantly, in all types of measures, the figure exhibits a substantial divergence of the frequencies of actual performance over the two *PRT* intervals located immediately below and above the target level. Presented in Figure 3, all histograms of the frequency distribution of *PRT* have a deep pit immediately below a target level (i.e., 100%) and, in contrast, a highest peak immediately above the target level. Specifically, actual performance's being at or slightly over a target level[12] is three to six times more likely than its being slightly below the target depending on the type of a performance measure. Such abnormality in the frequency distribution (i.e., a "discontinuity at zero") of actual performance relative to a benchmark is considered as evidence of a managers' strong incentive to achieve a target. However, the prior evidence exists, in most cases, for a specific type of performance benchmark—zero accounting earnings.

Distribution of Performance by Year. Panel C provides further description about the distribution of performance by year and performance measure type. Panel A shows a gentle rising trend of *PRT* and target achievability. The mean (median) of *PRT* increases from 98.7% (97.8%) to 131.4% (104.3%) during the sample period. The mild

[12] The bin width for "All measures" is 1% while those for the other types of measures are 2%.

increase of the median of *PRT* in 2009 compared with the sharp increase of the mean and the huge standard deviation suggest the likely presence of large outliers. Indeed, a greater number of huge outliers exist in 2009 compared to the other years.[13] We attribute it to greater difficulty of target setting after an economic shock.

Insert Table 3 about here.

Descriptive Evidence of Serial Correlations

Correlations. Table 3 presents correlations between key variables. It reports positive associations of *ex post* target achievability ($Achieved_t$) and the previous period's performance (PRT_{t-1}) and overachievement ($PRT_{t-1} \geq 120\%$) with the current period's performance to target (PRT_t). It also shows that the current period's target achievement ($Achieved_t$) is positively associated with the previous period's performance (PRT_{t-1}), achievement ($Achieved_{t-1}$) and overachievement ($PRT_{t-1} \geq 120\%$), and negatively associated with the previous period's underachievement ($PRT_{t-1} < 80\%$) and a measure's being classified as common (*Common*). The univariate evidence suggests that performance relative to a target is serially correlated. Following Indjejikian and Nanda (2002), we interpret a serial correlation as evidence of incomplete target revision.

Insert Table 4 about here.

Insert Figure 4 about here.
---

[13] Specifically, in 2009 there are 16 observations whose PRT is greater than 200%, while there are zero, seven, ten, and two in the other years. Further, the largest three outliers (1,272.2%, 1,103.0%, and 1,093.0%) in the year are almost twice and three times as large as the maximum PRT (627.3%) and the second largest (410.1%) of all the other years respectively. Excluding the 16 outliers, the mean and the median of the other 232 observations are 103.6 and 102.3.

Linear Relationships of PRTs between Years. Table 4 and Figure 4 illustrate the inter-temporal relationship of performance to a target between two consecutive years. The table is constructed to present the transition probability that is defined as the probability of transitioning from one *PRT* partition to another in a following period,[14] and Figure 4 is a graphical representation of Table 4. Each *PRT* partition other than the ones at both ends has a 10% *PRT* interval starting from every 10% point. For example, the "1.0~1.1" partition ranges from 100% to 110%. Each cell represents the transition probability. For example, the cell where the "1.0~1.1 PRT_{t-1}" row intersects with the "1.0~1.1 PRT_t" column is 43.6; this indicates that conditional on that PRT_{t-1} is between 100% and 110%, the probability of PRT_t ending up between 100% and 110% is 43.6%.

Serial correlations of performance between consecutive years would be manifested as high probabilities along the cells on the diagonal from the bottom left to the top right—colored in gray. However, the linear relationship seems not evident in the table. On one hand, it is true that some high probabilities are positioned along the diagonal. On the other hand, many other high probabilities lie off the diagonal line as well. For example, the maximum values in the "0.8~0.9" and "1.4~1.5" rows are 26.9% and 30.0% which are located off the diagonal (both in "1.0~1.1" column). Panel A of Figure 4 plots the observations in a graph and confirms the weak linear relationship. Specifically, the scatter plots are (1) populated between 80% and 100% on x- and y-axis; and (2) well scattered around in the other area, which raises a doubt as to, if any, serial correlations.

[14] It is a conditional probability denoted as $\Pr(PRT_t=j|PRT_{t-1}=i)$ where t represents time, and i and j represent a *PRT* partition which each period's *PRT* falls in.

Column "*Similar*" provides the probability of the current period's *PRT*'s (PRT_t) being similar to the previous period's *PRT* (PRT_{t-1}).[15] Again, a serial correlation of performance would have resulted in high probabilities along the antidiagonal and hence corresponding high values across all the cells in this column, which is not the case. Instead, the transition probabilities of having similar *PRT*'s between two years are high only in the middle partitions from 90% to 120%.[16] The finding suggests that a serial correlation of *PRT*'s is limited to the range. Column "*1.0~1.2*" adds up the "1.0~1.1" and "1.1~1.2" partitions where the current target is achieved but not overachieved. Interestingly, the transition probabilities to these two *PRT* partitions (i.e., 1.0~1.2) from any *PRT* partitions are generally higher than the probabilities of having similar *PRT*'s in two consecutive years. Lastly, Column "*Achieved*" is the probability of the current target being achieved conditional on PRT_{t-1}. The target achievability is significantly higher for previously achieved performance items (i.e., at or above PRT_{t-1} of 1.0) than for missed ones (i.e., below PRT_{t-1} of 1.0).[17]

Overall, the analysis of transition probabilities and related statistics in Table 4 provides important backgrounds for the main tests in the following sections. First, it shows that the likelihoods of target achievement in a measure between years are serially correlated and that the levels are also serially correlated but they are only within a certain range between 100% and 120%.[18] This suggests that the information may not be fully reflected in new targets at the performance level. Second, outside the *PRT* range (i.e.,

[15] We define similar performance as PRT_t being in any of three neighboring *PRT* partitions centering the previous' period's PRT partition: so to speak, $\Pr(PRT_t=j|PRT_{t-1}=i)$ for i and j such that $i-1 \leq j \leq i+1$.
[16] See Panel B of Figure 4 for its graphical presentation.
[17] The target achievability between the two groups (i.e., previously achieved vs. missed) is significantly different ($p<0.001$). The test result is not tabulated.
[18] Where targets are closely missed or achieved but not overachieved.

100%~120%), targets seem to be revised to incorporate the information about managers' productivity and/or to eliminate uncontrollable factors that are identified in the previous period's performance. For example, for previous poor performance, target achievability is reset so that new targets can be achieved with 50% of probability[19] rather than be achieved with similar probability as in the previous period.

Multivariate Analysis

Target Achievement

Table 5 presents the odd-ratios from logit regression models predicting the likelihood of target achievement. To estimate the likelihood of target achievement, we include 695 measure-years of which the *PRT* data are available for at least two consecutive years. The variables of our interest in the regression models are indicator variables of whether a performance target is achieved ($Achieved_{t-1}$), overachieved ($PRT_{t-1} \geq 120\%$), and underachieved ($PRT_{t-1} < 80\%$). To control for the variations due to the use of different types of performance measures, we include dummy variables that indicate the type of a performance measure (*Financial* and *Common*). We also include the importance weight of a performance measure (*Weight*). For all regressions, business entity dummies and year dummies are used as fixed effects to address cross-sectional and serial correlations. Model 1 includes three dummy variables of our interest without control variables, while Models 2 and 3 control for the type and the importance weight of a performance measure in different specifications.

Insert Table 5 about here.

---

[19] The target achievability of previously missed performance targets is not different from 50% (p=0.5430).

Table 5 reports that, in all models, the probability of a target's being achieved in a year is strongly associated with the target achievement in the previous year. The odds-ratio on $Achieved_{t-1}$ is significantly greater than one[20] in all models. The odds-ratios suggest that the odds of achieving a target are approximately 70% higher when its previous target was achieved than when it was missed. Accordingly, the findings suggest a serial correlation of target achievement between two years, supporting H1. On the other hand, the other two indicator variables for overachievement ($PRT_{t-1} \geq 120\%$) and underachievement ($PRT_{t-1} < 80\%$) are statistically insignificant. First, the insignificant $PRT_{t-1} \geq 120\%$ suggests that overachievement does not help to further improve the likelihood of achievement in the following period over the positive effect of previous target achievement. Second, with insignificant $PRT_{t-1} < 80\%$, we find no incremental effect of underachievement beyond the effect of missing a previous target. Thus, the findings in Table 5 support H3 and H4. With regards to the effects of performance characteristics, Models 2 and 3 show that performance targets in common measures are less likely to be achieved than those in unique measures.

Insert Table 6 about here.

Performance Relative to a Target

To test H1 to H6, we estimate performance to a target for performance measure *i* in year *t* ($PRT_{i,t}$) with its lagged value ($PRT_{i,t-1}$), indicators of PRT_{t-1} partitions ($Missed_{i,t-1}$, $PRT \geq 120\%_{i,t-1}$, and $PRT < 80\%_{i,t-1}$), indicators of the type of the performance measure ($Financial_{i,t}$ and $Common_{i,t}$), the measure's importance weight ($Weight_{i,t}$), and interaction

[20] Note that an odds-ratio greater (less) than one is equivalent to a positive (negative) coefficient in a logit regression.

variables of all these variables with $PRT_{i,t-1}$. We also include year and business entity dummies to control for cross-sectional correlations across business entities in each year and for business entity wide effects. The regression parameters are estimated using the Feasible Generalized Least Squares (FGLS) method to address heteroskedasticity and auto-correlations between observations. The model is expressed as follows:

$$\begin{aligned} PRT_{i,t} = {} & \beta_0 + \beta_1 PRT_{i,t-1} \\ & + \beta_2 Missed_{i,t-1} + \beta_3 PRT \geq 120\%_{i,t-1} + \beta_4 PRT < 80\%_{i,t-1} \\ & + \beta_5 Financial_{i,t} + \beta_6 Common_{i,t} + \beta_7 Weight_{i,t} \\ & + \beta_8 PRT_{i,t-1} Missed_{i,t-1} + \beta_9 PRT_{i,t} PRT \geq 120\%_{i,t-1} \\ & + \beta_{10} PRT_{i,t-1} PRT < 80\%_{i,t} + \beta_{11} PRT_{i,t-1} Financial_{i,t} \\ & + \beta_{12} PRT_{i,t-1} Common_{i,t} + \beta_{13} PRT_{i,t-1} Weight_{i,t} \\ & + year\ dummies + business\ entity\ fixed\ effects + \varepsilon. \end{aligned}$$

Table 6 presents three FGLS regression models in which *PRT* for a previous period is partitioned into two (achieved and missed, in Model 1), three (overachieved, achieved, and missed, in Model 2), and four (overachieved, achieved, missed, and underachieved, in Model 3) sections. In these regressions, our interests are on the coefficients of $PRT_{i,t-1}$ (β_1) and its interaction terms with the other variables (β_8 to β_{13}).

Test of H1

Analogous to Indjejikian and Nanda's (2002), if a new target is set with previous performance fully incorporated, target achievement levels should not be serially correlated, and thus β_1 should not be different from zero. Anticipating firms' commitment to incomplete incorporation of previous performance in setting new targets, we expect positive and significant β_1 to support H1.

In Model 1 that is similar to Indjejikian and Nanda's (2002) design with two PRT_{t-1} partitions—*achieved* and *missed*, the coefficient of PRT_{t-1} does not pick up statistical significance. As a potential cause of the difference, we note the findings in

Table 4 and Figure 4. They suggest that partitioning *PRT* simply into *achieved* and *missed* may not be descriptive of the non-linear characteristic of the inter-temporal relationship of *PRT*'s. Consistent with these observations from Table 4 and Figure 4, Models 2 and 3 with more *PRT* partitions have their β_1 positive and significant at 99% level. The findings suggest that *PRT* of a performance measure is, in general, positively associated with its previous *PRT*.[21] Therefore, the results in Table 6 support H1.

Test of H2

To test H2, we examine whether the total effect of $PRT_{i,t-1}$, or $\beta_1 + \beta_8$ (for $Missed_{i,t-1}$) differs significantly from zero. Given $\beta_1>0$, $\beta_1 + \beta_8$ may be zero, positive, or negative. First, $\beta_1 + \beta_8$ not different from zero indicates the absence of serial correlation for previously missed performance measures. This is consistent with what is expected in H2a—fair and neutral target revision, or complete incorporation of previous performance in new targets. Second, $\beta_1 + \beta_8$ greater than zero indicates that a positive serial correlation exists even for previously missed performance measures. This suggests that previous poor performance is penalized such that a new target is set as difficult as the previous one. Thus, a positive sum of the coefficients supports H2b—penalization of past poor performance, or no downward adjustment of target levels. Third, $\beta_1 + \beta_8$ less than zero indicates a negative serial correlation for those who failed to meet previous targets—rewarding poorer performance and penalizing better performance, which hardly satisfies our commonsensical criterion.

In Table 6, we find that the coefficient of the interaction between PRT_{t-1} and $Missed_{t-1}$ (β_8) is negative and significant at 99% level in Models 2 and 3. A negative β_8

[21] As long as performance targets are achieved but not overachieved (i.e., $100 \leq PRT_{t-1} < 120$) where $Missed_{t-1}=0$, $PRT_{t-1} \geq 120\%=0$, and $PRT_{t-1} < 80\%=0$.

indicates that a positive association of *PRT*'s between two years is mitigated to some extent. As discussed above, whether the positive association is fully eliminated can be tested with $\beta_1 + \beta_8$. The χ^2 test results for the coefficient sums ($\beta_1 + \beta_8$) reported at the bottom do not differ significantly from zero. The results suggest that KC adjusts targets to fully reflect past performance information on a fair and neutral basis—not compensating for, nor penalizing poor performance. This supports H2a as opposed to H2b.

Test of H3 and H4

H3 states that firms eliminate, if any, a luck factor or a previous period's excess *PRT* such that the overachievement has no incremental effect beyond what is expected from target achievement in the previous period. In Table 6, the zero interaction effect ($\beta_1 + \beta_9$) in Models 2 and 3 suggests that the positive and significant main effect of PRT_{t-1} is obliterated by the negative and significant interaction between PRT_{t-1} and *overachievement* ($PRT_{t-1} \geq 120\%$), which indicates that KC, indeed, revises targets to eliminate any luck factor and/or to correct its evaluation of managers' productivity that was previously underestimated. Thus, Table 6 provides no support for a positive serial correlation of *PRT*'s for previously overachieved performance measures.

Whereas H3 predicts firms' isolation of a luck factor, H4 tests firms' compensation for bad luck. To this end, H4 investigates the sum of three coefficients ($\beta_1 + \beta_8 + \beta_{10}$). We find that the coefficient sum in Model 3 is not significantly different from zero, which suggests that KC adjusts a target difficulty level downward to remove the impact of huge previous bad luck.

Test of H5 and H6

H5 and H6 investigate how different characteristics of performance measures included in bonus plans affect a firm's target revision behavior. We examine the coefficients of interaction terms to test incremental impacts of the characteristics of performance measures. First, as to the effect of financial performance measures, the coefficient of the interaction between PRT_{t-1} and $Financial_t$ (β_{11}) is positive and significant in all models. The finding suggests that a positive serial correlation of *PRT*'s for financial measures is intensified. In plain English, when a previous target is achieved, KC tends to reflect past performance of financial measures in a new target less completely as compared to when it does with that of nonfinancial measures. By doing so, KC may mitigate managers' concern about the KC's reneging, or ratcheting of financial targets. On the other hand, the positive coefficient also suggests that KC penalizes poor performance (i.e., missing a target) by letting prior poor performance likely to continue in the following period. Second, the coefficient of the interaction between PRT_{t-1} and $Common_t$ (β_{12}) is insignificant in all models. The findings suggest that, in setting new targets for common measures, KC makes commitment to incomplete use of previous performance information as much as it does for unique measures. Third, investigating the effect of importance weights, we find no evidence of their incremental effect on a serial correlation. This contrasts to our expectation, a positive β_{13}, that greater importance weights in bonus formula would increase managers' concern about undesirable target ratcheting.

In the model specification, we may well suspect multicollinearity between *Financial*, *Common*, and *Weight* as Table 3 reports significant correlations among the variables. To address potential multicollinearity, we run separate regressions with only

one indicator of these performance characteristics, among *Financial*, *Common*, and *Weight*, at one time. Untabulated results show that the coefficient of PRT_{t-1} interacted with *Financial* remains significant (β=0.146, p=0.011), that with *Common* becomes significant even at 95% level (β=0.085, p=0.026), and that with *Weight*, on the other hand, gets weaker (β=0.122, p=0.764).

To summarize, the results reported in Table 6 reject H5 (in a null-form) and H6. The positive and significant coefficients (β_{11} and β_{12}) indicate KC's different behavior in revising targets of financial and common measures. On one hand, when a target in financial or common measures was met, KC makes a stronger commitment to incomplete use of past performance for financial or common measures than for nonfinancial or unique measures. When a target was missed, KC, on the other hand, gives stronger penalty for past poor performance in financial and common measures—by setting new targets in these measures as difficult as previous targets—than poor performance in nonfinancial and unique measures. As to the effects of importance weights, we find that, contrary to our prediction, importance weights do not affect the firm's target revision behavior. Among these findings, it is notable that KC commits to incomplete use of past performance in target revision of financial or common measures despite their greater visibility and accordingly less information asymmetry. As a result, KC's such a commitment allows managers to achieve targets in financial and/or common performance measures more consistently (i.e., a high serial correlation) than the others.

VI. CONCLUSION

Firms often set performance targets based on information about a manager's productivity disclosed in past performance. Theoretically, setting a target based on the

same person's previous performance is an efficient way to identify his/her productivity (Milgrom and Roberts, 1992). Such a practice, however, may give rise to a serious incentive conflict, or the "ratchet effect," that discourages workers from continuing to make their best effort once a performance target is reached. As a remedy to the problem, prior literature suggests a firm's commitment to reflecting past performance in a new target to an incomplete extent so that managers get less concerned about potential backfiring of good past performance on a new target (Laffont and Tirole, 1993; Milgrom and Roberts, 1992).

Acquiring access to unique and exclusive data of annual performance evaluation for very high-ranking executives who manage companies or large business units in a large multinational conglomerate, we investigate the conglomerate's target revision practice. Specifically, we empirically test (1) whether a firm demonstrates a commitment to incomplete use of past performance in revising target levels; and (2) whether it reacts differently responding to different levels of previous target achievement and to measures in different types or with different importance weights. Along with the investigation of the main research question, we also examine managers' behavior in achieving targets in different types of performance measures.

The evidence from the frequency distributions of performance and descriptive statistics is generally evident and consistent with our expectation about managers' target-achieving behavior. First, we find a discontinuity at a target level in the frequency distributions of *PRT*. The discontinuity exists in the distributions of *PRT* irrespective of the types of performance measures, which suggests managers' strong incentives to achieve targets of any performance measures. Second, the analysis of transition

probability suggests a serial correlation of *PRT*'s of a performance measure between two consecutive years, especially within a certain range of target achievement level (i.e., *PRT* between 90% and 120%).

We find that the firm makes a credible commitment to incomplete incorporation of past performance into a new target. With multivariate analyses using logit regression and GLS regression techniques, we consistently find that the likelihood and the level of target achievement are serially correlated. In particular, a target specified in a performance measure is more likely to be achieved when its previous target was achieved. Moreover, the target achievement level in a period is positively associated with its previous achievement level with a certain previous *PRT* range. As suggested in prior literature, we interpret the positive serial correlation of *PRT*'s between years as the evidence of firms' commitment.

Investigating the effects of over- and under-achievement, we present evidence that the firm corrects, if any, a luck factor and its underestimation of a manger's productivity such that overachievement (underachievement) does not continue, or does not improve (harm) the likelihood and the level of target achievement in the following period. Further, our examination of the effects of performance measure type shows that *PRT*'s of financial and common measures show stronger consistency in time. This suggests that a firm may give a stronger reward (penalty) for past good (bad) performance in measures with greater visibility and comparability and less information asymmetry than it does for performance in other types.

Despite the interesting and evident findings, our study is subject to limitations. As in typical field studies, our data are from practically one large firm. The data are collected

for 133 business-entity-year observations from 29 corporations and large business units for five years. These business entities are, in effect, governed by one holding company, and thus under one organizational policy. For this reason, we find the firms' systematic and consistent target setting behavior. However, for the same reason, the generalizability of our findings to other contexts or to other firms may be limited. Further, although our hypotheses are developed based on prior literature, we cannot rule out alternative explanations that may better explain our findings including target revision behavior varying with the type of performance measures. Probably, field-based research components such as interviews or experimental research would help to build solid foundation and to add credibility. That being said, future research would consider how our findings from this research can be extended to other organizations or explore research methods to improve internal validity of our empirical findings.

REFERENCES

Anderson, S.W., Dekker, H.C., Sedatole, K.L., 2010. An empirical examination of goals and performance-to-goal following the introduction of an incentive bonus plan with participative goal setting. *Management Science* 56, 90-109.

Bouwens, J., Kroos, P., 2011. Target ratcheting and effort reduction. *Journal of Accounting and Economics* 51, 171-185.

Burgstahler, D., Chuk, E., 2012. What have we learned about earnings management? Correcting disinformation about discontinuities: Unpublished working paper.

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

Burgstahler, D., Eames, M., 2006. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. *Journal of Business Finance & Accounting* 33, 633-652.

Choi, B.H., Lee, K.T., 2004. An empirical study of the ratchet effect in performance evaluation system. *Korean Accounting Review* 29, 293-322.

Degeorge, F., Patel, J., Zeckhauser, R., 1999. Earnings management to exceed thresholds. *The Journal of Business* 72, 1-33.
Hayn, C., 1995. The information content of losses. *Journal of Accounting and Economics* 20, 125-153.
Holthausen, R.W., Larcker, D.F., Sloan, R.G., 1995. Annual bonus schemes and the manipulation of earnings. *Journal of Accounting and Economics* 19, 29-74.
Indjejikian, R., Nanda, D., 1999. Dynamic incentives and responsibility accounting. *Journal of Accounting and Economics* 27, 177-201.
Indjejikian, R., Nanda, D., 2003. Reply to: Dynamic incentives and responsibility accounting: A comment. *Journal of Accounting and Economics* 35, 437-441.
Indjejikian, R.J., Matějka, M., Merchant, K.A., Van der Stede, W.A., 2012. Earnings targets and annual bonus incentives: Unpublished working paper.
Indjejikian, R.J., Nanda, D., 2002. Executive target bonuses and what they imply about performance standards. *Accounting Review*, 793-819.
Ittner, C.D., Larcker, D.F., Rajan, M.V., 1997. The choice of performance measures in annual bonus contracts. *Accounting Review*, 231-255.
Kaplan, R.S., Norton, D.P., 1992. The balanced scorecard–measures that drive performance. *Harvard Business Review* 70, 71-79.
Kaplan, R.S., Norton, D.P., 1996. The balanced scorecard: Translating strategy into action: Harvard Business Press.
Laffont, J.J., Tirole, J., 1993. A theory of incentives in procurement and regulation: the MIT Press.
Lal, M., Dunk, A.S., Smith, G.D., 1996. The propensity of managers to create budgetary slack: A cross-national re-examination using random sampling. *The International Journal of Accounting* 31, 483-496.
Leone, A.J., Rock, S., 2002. Empirical tests of budget ratcheting and its effect on managers' discretionary accrual choices. *Journal of Accounting and Economics* 33, 43-67.
Libby, T., Lindsay, R.M., 2010. Beyond budgeting or budgeting reconsidered? A survey of north-american budgeting practice. *Management Accounting Research* 21, 56-75.
Libby, T., Salterio, S.E., Webb, A., 2004. The balanced scorecard: The effects of assurance and process accountability on managerial judgment. *The Accounting Review*, 1075-1094.
Lipe, M.G., Salterio, S.E., 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review*, 283-298.
Merchant, K.A., 1985. Budgeting and the propensity to create budgetary slack. *Accounting, Organizations and Society* 10, 201-210.
Merchant, K.A., Manzoni, J.F., 1989. The achievability of budget targets in profit centers: A field study. *The Accounting Review*, 539-558.
Merchant, K.A., Van Der Stede, W.A., 2007. Management control systems: Performance measurement, evaluation and incentives: Financial Times/Prentice Hall.
Milgrom, P., Roberts, J., 1992. Economics, organization and management: Prentice-Hall Englewood Cliffs, NJ.
Murphy, K.J., 2001. Performance standards in incentive contracts. *Journal of Accounting and Economics* 30, 245-278.

Slovic, P., MacPhillamy, D., 1974. Dimensional commensurability and cue utilization in comparative judgment. *Organizational Behavior and Human Performance* 11, 172-194.

Umapathy, S., 1987. Current budgeting practices in u.S. Industry: The state of the art: Quorum Books.

Weitzman, M.L., 1980. The" ratchet principle" and performance incentives. *The Bell Journal of Economics*, 302-308.

Figure 1
Timeline of KC's Budgeting, Target-setting, and Performance Evaluation Process

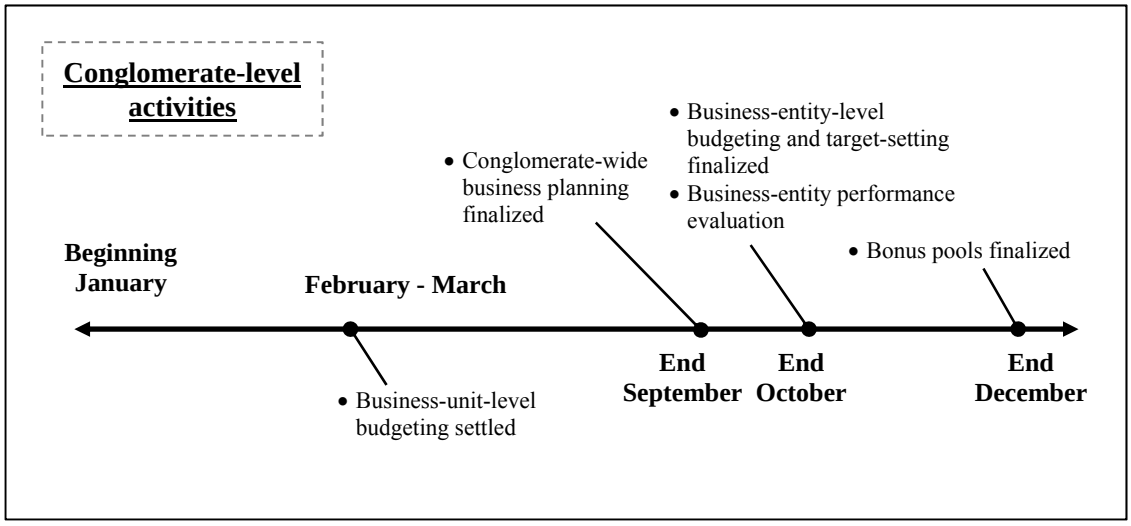


Figure 2
KC's Short-Term Incentive Plan

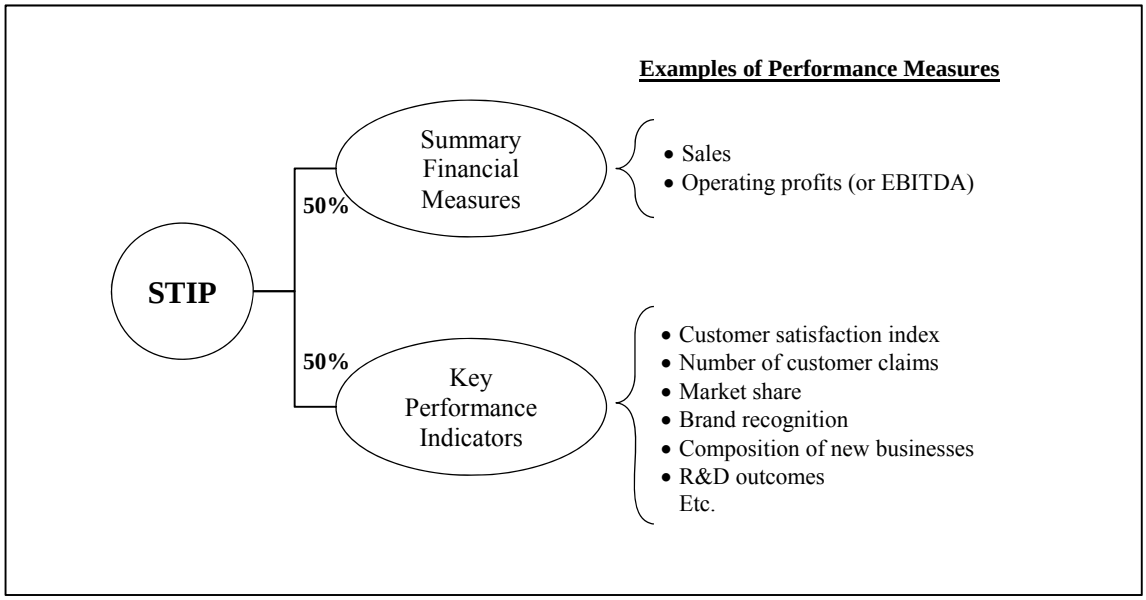
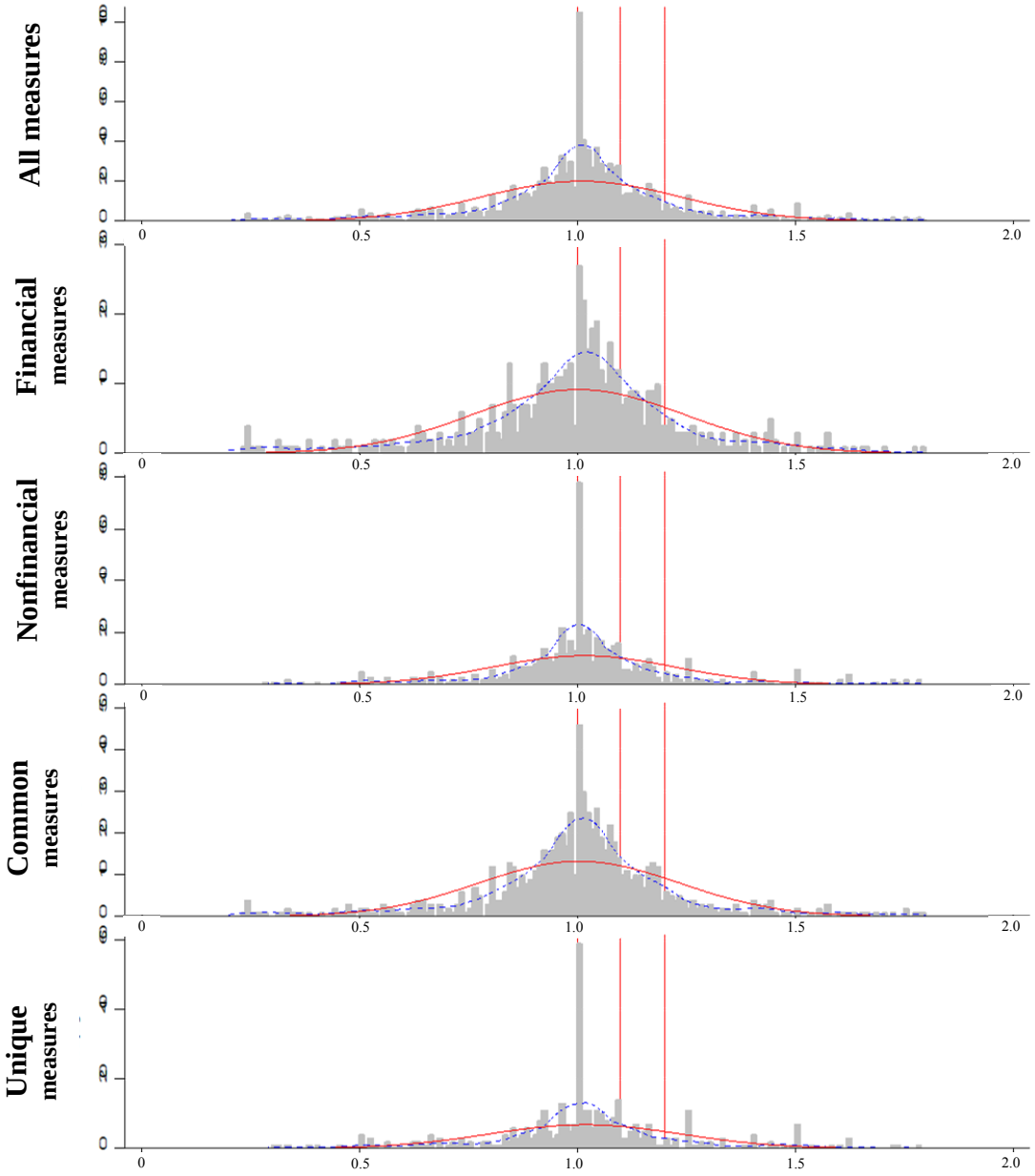


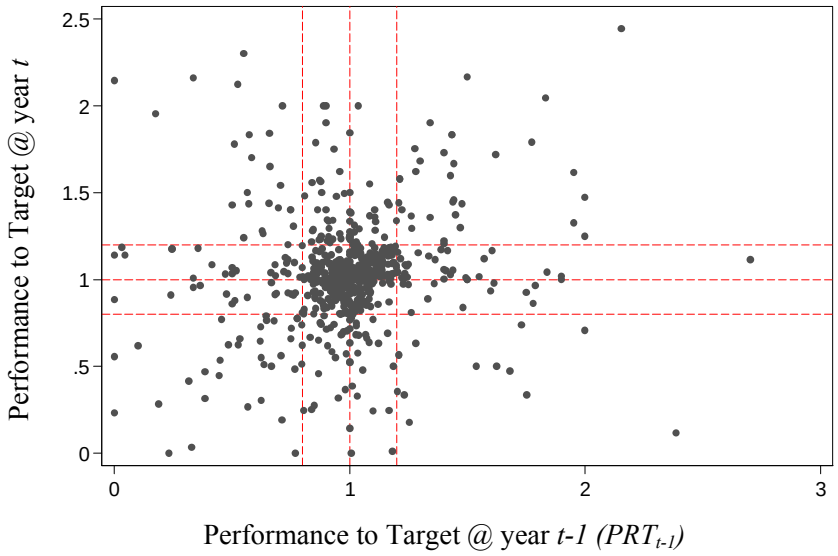
Figure 3
Distribution of Performance Relative to Target



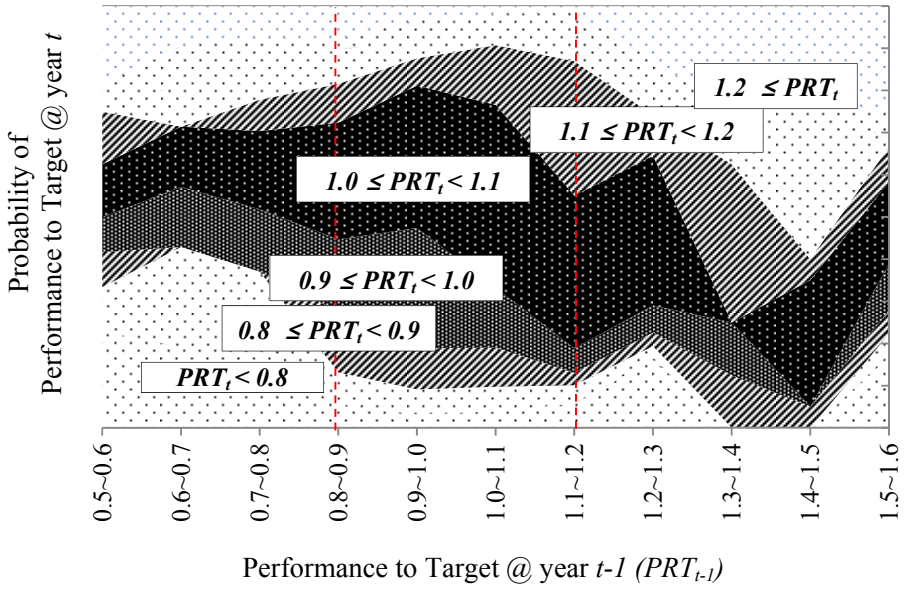
Solid lines indicate probability density functions of a normal distribution with the mean and the standard deviation of each performance measure group, while dashed lines represent kernel density estimation. The bin width for "All measures" is 0.01 (i.e., 1%) of performance relative to target (PRT) while those for the others are 0.02 (2%). Vertical auxiliary lines indicate 1 (100%), 1.1 (110%), and 1.2 (120%) of PRT respectively.

Figure 4
Performance to Target between Two Consecutive Years

Panel A: Scatterplot



Panel B: Stacked Area



In Panel A, auxiliary dashed lines indicate 0.8 (80%), 1 (100%), and 1.1 (110%) of performance relative to target from the bottom (the left) respectively.

Table 1
Performance Measure Classifications

Panel A: Financial vs. Non-Financial

	Types	2005	2006	2007	2008	2009	Total
Financial	Profit	48	50	46	53	58	255
	Sales	26	26	26	38	47	163
	Composition	24	23	24	23	17	111
	Business initiatives	6	5	9	8	7	35
	Cost management	5	4	6	2	2	19
	R&D related financials	4	1	1	3	2	11
	Company specific	1	1	3	-	2	7
	Financial	**114**	**110**	**115**	**127**	**135**	**601**
Non-Financial	KC employee survey	27	25	27	27	26	132
	Customer satisfaction	18	15	17	18	18	86
	Market share	9	12	11	15	15	62
	HRM indexes	5	14	13	2	1	35
	Brand recognition	7	7	8	5	4	31
	R&D related indexes	5	5	5	3	4	22
	Business initiatives	-	1	1	1	-	3
	Company specific	44	60	56	31	45	236
	Non-Financial	**115**	**139**	**138**	**102**	**113**	**607**
Total		**229**	**249**	**253**	**229**	**248**	**1,208**

Panel B: Common vs. Unique

	Types	2005	2006	2007	2008	2009	Total
Common	Profit	48	50	46	53	58	255
	Sales	26	26	26	38	47	163
	Composition	24	23	24	23	17	111
	KC employee survey	27	25	27	27	26	132
	Customer satisfaction	18	15	17	18	18	86
	Market share	9	12	11	15	15	**62** [a]
	Common	**152**	**151**	**151**	**174**	**181**	**809**
Unique	Business initiatives	6	6	10	9	7	38
	HRM indexes	5	14	13	2	1	35
	R&D	9	6	6	6	6	33
	Brand recognition	7	7	8	5	4	31
	Cost management	5	4	6	2	2	19
	Company specific	45	61	59	31	47	243
	Unique	**77**	**98**	**102**	**55**	**67**	**399**
Total		**229**	**249**	**253**	**229**	**248**	**1,208**

a. The median frequency of adoption (computed without 243 company specific measures) is 62, which constitutes the benchmark separating the common vs. unique performance measure types.

Table 2
Descriptive Statistics

Panel A: Use of performance measures

	Year					
	2005	2006	2007	2008	2009	Total
No. of measures	229	249	253	229	248	1,208
Business entity-year	27	27	27	26	29	133
Measures per business entity	8.5	9.2	9.4	8.8	8.6	9.1

Panel B: Performance measures used in Short-Term Incentive Plan (STIP)

	N	Mean	Std Dev.	1Q	Median	3Q
Performance measures used in STIP						
Used for one year [a]	237					
Used for at least two consecutive years [b]	971					
Financial measures	601					
Nonfinancial measures	607					
Common measures	809					
Unique measures	399					
Importance weight	1,208	10.8	6.1	5.0	10.0	12.5
Performance						
Target achieved	1,208	59.1				
Target overachieved (*PRT* [c] greater than 120%)	1,208	15.7				
Target missed by far (*PRT* [c] less than 80%)	1,208	12.4				
Performance relative to target (in *PRT* [c])	1,208	108.6	71.3	92.0	101.1	112.0

Panel C: Performance Relative to Target by Year

Year	N	Mean	Std Dev.	1Q	Median	3Q	>100 [d]	>120 [e]	<80 [f]
2005	229	98.7	12.8	92.6	97.8	105.6	45.9	3.9	3.9
2006	249	101.0	48.7	85.2	100.0	109.5	54.6	14.9	20.5
2007	253	105.5	40.8	89.6	101.6	117.5	60.9	22.5	17.0
2008	229	105.5	31.7	94.9	102.1	115.4	62.4	16.6	9.6
2009	248	131.4	137.9	97.6	104.3	115.5	70.6	19.8	10.1
Full Sample	1,208	108.6	71.3	92.0	101.1	112.0	59.1	15.7	12.4

Units in Panels B and C are per cent except for the number of observations (i.e., N).
a. The number of performance measures that are adopted in STIP for at least two consecutive years.
b. The number of performance measures that are adopted in STIP for a year or not adopted for consecutive years.
c. *PRT* is for *Performance Relative to Target.*
d. The proportion of performance targets that are achieved (i.e., the performance-relative-to-target equals or exceeds 100.).

e. The proportion of performance targets that are overachieved (i.e., the performance-relative-to-target equals or exceeds 120.).
f. The proportion of performance targets that are missed by a considerable margin (i.e., the performance-relative-to-target is less than 80.).

Table 3
Correlations

	1	2	3	4	5	6	7	8
1. Performance to target *(PRT_t)*								
2. *Achieved* $_t$	0.330***							
3. PRT_{t-1}	0.130***	0.083**						
4. *Achieved* $_{t-1}$	0.050	0.196***	0.507***					
5. $PRT_{t-1} \geq 120$	0.092**	0.077**	0.601***	0.368***				
6. $PRT_{t-1} < 80$	-0.048	-0.114***	-0.495***	-0.424***	-0.156***			
7. *Financial*	0.048*	-0.024	-0.076**	-0.004	-0.055	0.050		
8. *Common*	0.015	-0.069**	-0.076**	-0.070*	-0.047	0.015	0.445***	
9. *Weight*	-0.010	-0.030	-0.044	-0.016	-0.046	0.008	0.531***	0.305***

*, **, *** indicate significance at the 10, 5, and 1 two-tailed confidence level respectively.

PRT_{t-1} is the actual performance relative to target at year t-1. ***Achieved***$_t$ is an indicator variable; 1 if the target of a performance measure is achieved in year *t* and 0 otherwise. ***$PRT_{t-1} \geq 120$*** is an indicator variable; 1 if the target of a performance measure is achieved with PRT greater than 120 in the previous year and 0 otherwise. ***$PRT_{t-1} < 80$*** is an indicator variable; 1 if the target of a performance measure is underachieved with PRT less than 80 in the previous year and 0 otherwise. ***Financial*** is an indicator variable whose value is assigned 1 if a performance measure is a financial performance measure and 0 otherwise. ***Common*** is an indicator variable whose value is assigned 0 if a performance measure is used to define a company specific target or it is adopted more frequently than others (i.e., the number of adoption is greater than its median), and 0 otherwise. ***Weight*** is the importance weight of a performance measure.

Table 4
Transition Probability

PRT @ t-1	Performance relative to target @ t: ≤0.6	0.6~0.7	0.7~0.8	0.8~0.9	0.9~1.0	**1.0~1.1**	1.1~1.2	1.2~1.3	1.3~1.4	1.4~1.5	≥1.5	[a] Similar	[b] 1.0~1.2	[c] Achieved
>1.5	16.7		5.6	8.3	11.1	19.4	5.6	2.8	2.8	2.8	25.0	27.8	25.0	58.3
1.4~1.5				5.0		30.0	5.0	10.0	10.0	15.0	25.0	50.0	35.0	95.0
1.3~1.4				12.5	12.5		37.5		12.5		25.0	12.5	37.5	75.0
1.2~1.3	16.7			3.3	6.7	36.7	10.0	3.3	3.3	6.7	13.3	16.7	46.7	73.3
1.1~1.2	5.9	2.9	1.5	2.9	7.4	36.8	29.4	2.9	4.4	4.4	1.5	69.1	66.2	79.4
1.0~1.1	3.9	2.9	2.5	9.3	14.2	43.6	14.2	4.9	2.5		2.0	72.1	57.8	67.2
0.9~1.0	3.3	2.6	3.3	9.8	28.1	34.0	6.5	5.2	2.0	2.0	3.3	71.9	40.5	52.9
0.8~0.9	5.1	5.1	3.8	14.1	17.9	26.9	10.3	6.4		2.6	7.7	35.9	37.2	53.8
0.7~0.8	18.5	3.7	14.8		14.8	18.5	7.4	3.7	3.7	7.4	7.4	18.5	25.9	48.1
0.6~0.7	19.0	4.8	19.0		14.3	14.3		9.5		9.5	9.5	42.9	14.3	42.9
≤0.6	22.9	8.3	2.1	8.3	8.3	12.5	12.5	2.1		4.2	18.8	31.3	25.0	50.0
No obs.@t-1 [d]	4.1%	2.1	3.7	11.5	23.1	30.7	10.1	5.2	1.2	1.6	6.8		40.78	55.53
Total	6.0	2.7	3.6	9.6	18.9	31.7	11.3	5.0	1.9	2.3	7.0		42.96	59.11

Units are per cent.

a. The proportion of performance measure items of which the performance at year *t* is achieved at a similar level as year *t-1*. We define similar performance as PRT_t being in any of three neighboring *PRT* partitions centering the previous' period's PRT partition—cells in gray: so to speak, $\Pr(PRT_t=j|PRT_{t-1}=i)$ for *i* and *j* such that $i-1 \leq j \leq i+1$.

b. The proportion of performance measure items of which the performance at year *t* is between 100 and 120 (i.e., Probability($100 \leq PRT_t < 120|PRT_{t-1}$)).

c. The proportion of performance measure items of which the performance at year *t* exceeds a target level (i.e., Probability($PRT_t \geq 100|PRT_{t-1}$)).

d. The proportion of performance measure items of which the performance at year *t-1* is not available.

Table 5
Logit Regression Models Predicting Target Achievements in Year *t*

Variable	Hypo (Prediction)	**Model 1** Odds Ratio (Std Err.)	**Model 2** Odds Ratio (Std Err.)	**Model 3** Odds Ratio (Std Err.)
$Achieved_{t-1}$	H1(>1)	1.755***	1.708**	1.696**
		(0.214)	(0.211)	(0.212)
$PRT_{t-1} \geq 120$	H3(=1)	0.872	0.842	0.867
		(0.277)	(0.273)	(0.274)
$PRT_{t-1} < 80$	H4(=1)	0.705	0.697	0.696
		(0.287)	(0.295)	(0.293)
Financial	H5(?)		1.140	
			(0.232)	
Common	H5(?)		0.552**	
			(0.236)	
Financial and common	H5(?)			0.575**
				(0.276)
Financial and unique	H5(?)			0.317**
				(0.471)
Nonfinancial and common	H5(?)			0.370***
				(0.263)
Weight	H6(>1)		0.360	0.156
			(1.574)	(1.587)
Constant		0.963	1.823	2.777*
		(0.426)	(0.500)	(0.553)
Business entity dummies		Yes	Yes	Yes
Year dummies		Yes	Yes	Yes
Probability > χ^2		0.000	0.000	0.000
Pseudo R^2		0.115	0.124	0.134
Observations		693	693	693

*, **, *** indicate significance at the 10, 5, and 1 two-tailed confidence level respectively. T-statistics provided in parentheses are based on clustered robust standard errors that account for heteroskedasticity and autocorrelation.

$Achieved_{t-1}$ is an indicator variable; 1 if the target of a performance measure is achieved in the previous year and 0 otherwise. ***$PRT_{t-1} \geq 120$*** is an indicator variable; 1 if the target of a performance measure is achieved with PRT greater than 120 in the previous year and 0 otherwise. ***$PRT_{t-1} < 80$*** is an indicator variable; 1 if the target of a performance measure is underachieved with PRT less than 80 in the previous year and 0 otherwise. ***Financial*** is an indicator variable whose value is assigned 1 if a performance measure is a financial performance measure and 0 otherwise. ***Common*** is an indicator variable whose value is assigned 0 if a performance measure is used to define a company specific target or it is adopted more frequently than others (i.e., the number of adoption is greater than its median), and 0 otherwise. ***Weight*** is the importance weight of a performance measure.

Table 6
GLS Regression Models Predicting Performance Levels in Year *t*

Variable	Hypo (Prediction)	**Model 1** Coeff. (Std Err.)	**Model 2** Coeff. (Std Err.)	**Model 3** Coeff. (Std Err.)
PRT_{t-1}	H1 (+)	0.065	1.321***	1.376***
(β_1)		(1.070)	(19.182)	(20.415)
$Missed_{t-1}$		0.032	1.336***	1.442***
(β_2)		(0.815)	(16.088)	(17.876)
$PRT_{t-1} \geq 120$			1.444***	1.464***
(β_3)			(18.704)	(17.127)
$PRT_{t-1} < 80$				-0.144
(β_4)				(-1.436)
Financial		-0.281***	-0.182**	-0.198**
(β_5)		(-3.181)	(-2.282)	(-2.339)
Common		-0.112	-0.013	-0.001
(β_6)		(-1.813)	(-0.218)	(-0.014)
Weight		0.834	0.081	0.455
(β_7)		(1.440)	(0.147)	(0.852)
PRT_{t-1}*$Missed_{t-1}$	H2 (−/0)	-0.034	-1.266***	-1.380***
(β_8)		(-0.851)	(-15.183)	(-17.249)
PRT_{t-1}*($PRT_{t-1} \geq 120$)	H3 (−)		-1.349***	-1.365***
(β_9)			(-19.723)	(-18.420)
PRT_{t-1}*($PRT_{t-1} < 80$)	H4 (0)			0.189
(β_{10})				(1.361)
PRT_{t-1}**Financial*	H5 (?)	0.330***	0.228***	0.246***
(β_{11})		(3.830)	(2.992)	(3.020)
PRT_{t-1}**Common*	H5 (?)	0.088	-0.030	-0.041
(β_{12})		(1.449)	(-0.523)	(-0.710)
PRT_{t-1}**Weight*	H6 (+)	-1.207**	-0.410	-0.794
(β_{13})		(-2.085)	(-0.756)	(-1.491)
Constant		1.004***	-0.306***	-0.367***
Business entity dummies		Yes	Yes	Yes
Year dummies		Yes	Yes	Yes
Observations		573	573	573
χ^2 tests: The effect of PRT_{t-1}				
$\beta_1 + \beta_8 = 0$	H2 (0/+)	0.032	0.055	-0.004
(*Missed*)		(0.6225)	(0.4441)	(0.9633)
$\beta_1 + \beta_9 = 0$	H3 (0)		-0.028	0.011
($PRT_{t-1} \geq 120$)			(0.6206)	(0.8474)
$\beta_1 + \beta_8 + \beta_{10} = 0$	H4 (0)			0.185
($PRT_{t-1} < 80$)				(0.241)

*, **, *** indicate significance at the 10, 5, and 1 two-tailed confidence level respectively. The coefficient estimates and standard errors in Table 6 are estimated with the Feasible Generalized Least Squares (FGLS) method that addresses heteroskedasticity and auto-correlation between the panel observations. ***PRT***$_{t-1}$ is the actual performance relative to target at year t-1. ***Missed***$_{t-1}$ is an indicator variable; 1 if the target of a performance measure is not achieved in the previous year (i.e., year *t-1*) and 0 otherwise. ***PRT***

$_{t-1}\geq$***120*** is an indicator variable; 1 if the target of a performance measure is achieved with PRT greater than 120 in the previous year and 0 otherwise. $\boldsymbol{PRT_{t-1}<80}$ is an indicator variable; 1 if the target of a performance measure is underachieved with PRT less than 80 in the previous year and 0 otherwise. ***Financial*** is an indicator variable whose value is assigned 1 if a performance measure is a financial performance measure and 0 otherwise. ***Common*** is an indicator variable whose value is assigned 0 if a performance measure is used to define a company specific target or it is adopted more frequently than others (i.e., the number of adoption is greater than its median), and 0 otherwise. ***Weight*** is the importance weight of a performance measure.