Target Ratcheting in Common and Unique Performance Measures

Citation for published version (APA):
Kim, J., Choi, B. H., & Merchant, K. A. (Accepted/In press). *Target Ratcheting in Common and Unique Performance Measures*. Paper presented at American Accounting Association - Management Accounting Section Midyear Meeting, Scottsdale, United States.

Target Ratcheting in Common and Unique Performance Measures

BYUNG HYUN CHOI
LG Economic Research Institute

JONGHWAN (SIMON) KIM*
University of Manchester

KENNETH A. MERCHANT
University of Southern California

This Version: DECEMBER 2017

* Corresponding author: Department of Accounting and Finance, Alliance Manchester Business School, University of Manchester, Booth St. E., Manchester M13 9SS, U.K. E-mail: jonghwan.kim@manchester.ac.uk.

We gratefully acknowledge the support of the company that provided our research data which must, unfortunately, go unnamed. We thank Michal Matejka, Tony Davila, Jan Bouwens, Stanley Baiman, Sander Van Triest, Martin Holzhacker, Kyung-Tae Lee, Dae-Hee Yoon, and other participants at 2013 AAA-Management Accounting Section Meeting, 2014 Manufacturing Accounting Conference, 2014 New Directions for Management Accounting, and research seminars at Bocconi University, Yonsei University, and the University of Technology Sydney for the helpful comments. Previous versions of this paper were circulated under the title "Firms' Commitment to Mitigate the Ratchet Effect under Annual Bonus Plans with Multiple Performance Measures."

Target Ratcheting in Common and Unique Performance Measures

ABSTRACT

While most performance measurement systems use multiple measures, research on target revision in multi-measure systems is rare. Examining the target-setting practice in such a system, we hypothesize that the performance variances in common and unique measures are differentially informative of an agent's effort level and this, in turn, affects a principal's decision on the degree of commitment to incomplete incorporation of past performance information into new targets. We analyze proprietary data on the performance evaluation of strategic unit managers in a Korean conglomerate to report (1) intertemporal serial correlations of target achievement in both types of measures and (2) partial (full) target difficulty revision in common (unique) measures. With the findings from outside an experimental setting, we contribute to the accounting literature on a common measures bias as well as on target ratcheting.

Keywords: *performance target, performance measures, target difficulty, target revision, ratchet effect*

Data Availability: *Data used in this study are provided by a proprietary source and cannot be made public.*

1. Introduction

In this paper, we examine a firm's target revision practice[1] in a multi-measure performance evaluation system. Setting a challenging but achievable target without losing motivational effects is essential for a successful management control system (Merchant et al. 2012). For such an efficient target, firms may rely partly or entirely on an agent's past performance (Milgrom and Roberts 1992) that is often measured and evaluated with more than a single measure (e.g., Bain & Company 2015; Holmström 1979; Kaplan and Norton 1992). Nevertheless, the theory and evidence of target-setting practice under multiple-measure settings are lacking. To fill the void, we develop our predictions based on two streams of literature–performance measure properties (e.g., Gibbs et al. 2009; Lipe and Salterio 2000) and target ratcheting (e.g., Indjejikian, Matějka, Merchant, et al. 2014; Indjejikian, Matějka, and Schloetzer 2014)–and investigate whether and how the past performance information in common and unique measures affect a principal's target revision decisions.

To address our research questions, we analyze a proprietary dataset containing the annual performance evaluation of strategic business unit managers in a large Korean conglomerate. The dataset consists of 1,208 performance-measure-year observations for 133 business unit-years from 2005 to 2009. Analyzing the intertemporal target revision in common and unique performance measures, we find that the asymmetric target ratcheting that rewards past target achievement with more achievable targets in the following period occurs for both types of measures, but the marginal units of performance variances that represent business unit managers'

[1] Target revision refers to a target setting practice by which targets are set as an incremental of those in the previous period based on the principal's evaluation of an agent's performance. It should be distinguished from target adjustment made in order to correct a substantial error in forecasts or to adjust for unexpected and uncontrollable shocks during or at the end of the same performance period.

effort exertion in the two types are used differently in target revision. Specifically, our results show that a positive performance variance in common measures is appreciated and renders the principal's commitment to partial ratcheting that allows part of the previous positive variance to continue in the subsequent period, while that in unique measures is disregarded and leads to near-full target ratcheting, eliminating most, if not the entire part, of the previous favorable variance. Interestingly, the stylized target revision pattern exists only within a certain performance interval, specifically within ±20% relative to targets, where approximately 70% of observations populate and that generally conforms to our common-sensible criterion for a normal performance.

With the novel findings about the differential use of measures of different properties in a performance evaluation and target setting context, this study contributes to the literature on performance measure properties (e.g., Gibbs et al. 2009; Lipe and Salterio 2000) and target ratcheting (e.g., Aranda et al. 2014; Indjejikian and Nanda 2002) in several ways. First, our access to a unique dataset allows us a direct and comprehensive look at a firm's target revision practice. The dataset contains the targets and actual performance for the complete set of performance measures that are included in the firm's annual bonus plan for its strategic business unit managers. Without such data, prior studies in this area instead depend on inferred performance (e.g., bonus payout in Indjejikian and Nanda, 2002) or survey responses (Indjejikian, Matějka, Merchant, et al. 2014; Indjejikian and Matějka 2006; Mahlendorf et al. 2014). Unlike prior studies, our study uses objective data about performance in different types of measures to report the consistent results and additional others.

Second, thanks to the data opportunity, this study makes the first investigation of target-setting practices in a multi-measure environment. Our dataset is structured in a hierarchy where multiple performance measures are observed for each business unit (manager) by year. It contrasts to the previous studies' one-on-one matching between a manager and a measure. In particular, the prior research focused solely on target revision for a single financial measure such as profitability (Bol and Lill 2015), earnings (Indjejikian, Matějka, Merchant, et al. 2014), and sales (Aranda et al. 2014) of which the properties are relatively homogeneous as compared to those for a set of other vastly heterogeneous types of measures in multi-measure systems.

Third, this study responds to Indjejikian et al.'s (2014) call for research on the effect of information asymmetry on the association between previous performance and revised targets. We add the consistent evidence of asymmetric target revision and partial ratcheting in common measures, while we also document the lack of a principal's commitment to incomplete use of past performance information (i.e., full ratcheting) for new targets in unique measures. With the findings, our study introduces a new moderator associated with information asymmetry and contributes to the recent development in the literature that discusses a moderating effect of relative performance information on target revision (Aranda et al. 2014; Bol and Lill 2015).

Lastly, our finding of a common measures bias in the decisions made by top management team in a large conglomerate is the first documentation of the phenomenon outside experimental labs and with the real business actors other than undergraduate or MBA students (e.g., Banker et al. 2004; Dilla and Steinbart 2005; Libby et al. 2004; Lipe and Salterio 2000). This makes the study more relevant.

In the following section, we review the 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.

2. Background and Hypothesis Development

When past performance is used as a benchmark for target-setting, new targets are often ratcheted up for good past performance (Leone and Rock 2002; Weitzman 1980; Milgrom and Roberts 1992). In this setting, agents may well perceive an upward target revision as a penalty and stop exerting further effort to avoid such an undue penalty for stretched effort exertion, which is an adverse incentive problem known as the "ratchet effect" (Bouwens and Kroos 2011; Laffont and Tirole 1988; Leone and Rock 2002; Murphy 2000). To address the problem, theory suggests that firms should not fully incorporate information about a manager's productivity revealed in past period performance into the next period's target (Indjejikian and Nanda 2003; Laffont and Tirole 1993; Milgrom and Roberts 1992; Indjejikian and Nanda 1999). Consistent with the theory, empirical studies have documented the evidence of a principal's commitment to or implicit agreement for, incomplete incorporation of past performance (Indjejikian and Nanda 2002; Indjejikian and Matějka 2006; Aranda et al. 2014; Bol and Lill 2015).[2]

Importantly, in such a dynamic incentive situation, a reliable estimate for an agent's productivity and effort is the key to developing an efficient level of commitment to partial target ratcheting at which agents are fully motivated for stretched targets yet with little ratchet effect (Baron and Besanko 1984; Indjejikian, Matějka, and Schloetzer 2014). Past performance, however, is an imperfect proxy because it contains measurement errors and exogenous signals that are little attributable to an agent's effort (Banker and Datar 1989; Feltham and Xie 1994)

[2] For a full review of the target ratcheting and adverse incentives literature, see Indjejikian, Matějka, and Schloetzer (2014).

and it is prone to an agent's deliberate manipulations (Demski 1998; Holthausen et al. 1995; Ittner et al. 1997; Smith 2002). The noise or information asymmetry components hinder the principal from evaluating an agent's productivity level accurately and accordingly inhibit an optimal decision on a new target (e.g., Chow et al. 2012; Guidry et al. 1999). Further, in a multi-measure performance measurement system, measures of different properties involve different degrees of information asymmetry. To the extent, a principal may find past performance information in some measures more useful (i.e., more informative of an agent's effort) than others. This may affect the principal's target setting behaviors.

Discussing performance measure properties, prior literature has supplied theories and evidence for the differential informativeness of performance measures in the contexts of incentive contracting and performance measurement system design. The literature shows consistently that performance measure properties are key inputs in performance measurement system design and some properties make some measures more valuable to principals than others in estimating agents' productivity (Gibbs et al. 2009). In particular, the classical agency theory models show that, given the multiple signals about agents, principals place greater weightings on the information in performance measures with greater reliability: specifically, greater precision or less noise (Banker and Datar 1989; Holmström 1979; Ittner et al. 1997), greater controllable risks (Baker 2002; Prendergast 2002; Raith 2008), greater congruity or less distortion (Bouwens and Lent 2006; Feltham and Xie 1994), and less manipulability (Courty and Marschke 2004; Demski et al. 2004; Healy 1985; Holthausen et al. 1995).

The literature, however, examines only exogenous performance measure properties that determine the principals' decisions over the choice of performance measures in the design of such systems (i.e., weightings on measures). Differently motivated, a few recent studies recognize some principal-side characteristics and contingencies to explain the principals' behavior in using the information in performance measures which are readily included in the measurement system, (e.g., Lipe and Salterio 2000). Lipe and Salterio (2000) bring the idea of a common measures bias from Slovic and MacPhillamy (1974) to argue that principals are limitedly cognitive understanding the information in unique performance measures and thus adopt simple heuristics (Payne et al. 1993) to ignore it. Consistent with the theory, Lipe and Salterio (2000) find that, evaluating performance in common and unique measures, their experiment participants use only common measures and disregard unique measures. Following the introduction of a principal's cognitive limitation in performance evaluation to the literature, several accounting studies have contributed to the line of literature, reporting the bias-mitigating effects of, for example, principals' experience and knowledge (Dilla and Steinbart 2005), provision of strategy information (Banker et al. 2004; Humphreys and Trotman 2011), and provision of assurance and accountability requirement (Libby et al. 2004).

In sum, the literature suggests that designing or operating performance measurement systems with multiple measures, principals value more (the information content in) the measures with less uncontrollable risk, distortion, and manipulability and with greater agent-controllable risk (Gibbs et al. 2009) and those demanding less of a principal's cognitive effort in the process of distinguishing relevant information and the other in a performance measure (Lipe and Salterio 2000), and vice versa. We attend to some of the properties that are associated with information asymmetry between a principal and agents that may disturb the principal's efficient decision-making over performance evaluation and target setting. Among the aforementioned properties, uncontrollable risk is, by definition, beyond both principals' and agents' control, and distortion of incentives is induced by the (inefficient) selection of performance measures. They hardly affect the information environments. To the contrary, controllable risk, manipulability, and cognitive effort have something to do with agents' or principals' deliberate choice of actions surrounding the information asymmetry between the two parties over the course of their use in a system. These properties are associated with both parties' *post-design* choices and shape different information environments.

In particular, agents would be willing to exploit their informational advantage, withholding the information about their ability to respond to controllable risks and attempting to manipulate performance measures whenever possible (Gibbs et al. 2009). Thus, the availability and amount of specific knowledge on their jobs and tasks (Raith 2008; Prendergast 2002) determine the degree of informational advantage. In this paper, we note that common measures can help principals to mitigate the agent's opportunistic, dysfunctional exploitation of information asymmetry and accordingly improve the reliability of information in the measures. It is because some generic performance measures such as return on assets, operating profits, and customer satisfaction are more frequently adopted across business units and consistently over time (Kaplan and Norton 1996; Lipe and Salterio 2000), and the common measures often make it easier for a principal to compare performance among agents and filter out common transitory changes from others (Banker et al. 2004; Lipe and Salterio 2000). Further, as a principal participates in the evaluation or discussion of the performance in a measure more often, she gets more knowledgeable of the measure. In the context of performance target setting using past information, such a continuous learning process improves the principals' ability to estimate agents' productivity. Moreover, evaluating the past performance in common measures is less demanding than doing so using unique measures (Lipe and Salterio 2000; Slovic and MacPhillamy 1974).

The above discussion suggests that common measures, in general, involve relatively less information asymmetry than unique measures. The differential informativeness or reliability of information content between the two types renders a principal's systematic target revision behaviors to honor the information in the former and to disregard that in the latter. This leads to a reasonable prediction that the principal's commitment to partial ratcheting is meaningful only when she has a reliable estimate of an agent's productivity and she likely (rarely) comes across the reliable estimate when using common (unique) measures. Developing testable hypotheses based on this idea, we attend to two outcome variables: target achievability and target difficulty revision. First, we examine how a principal sets the likelihood of target achievement observing past performance variances. Second, we view performance variances as *ex post* target difficulty measures and compare them between years to examine a principal's revision of the target difficulty level.

2.1 Revision of Target Achievability

In a dynamic incentive context, a principal is keen to keep the agent's continuing effort and therefore commits to partial ratcheting of a target. The target ratcheting literature has consistently documented the principal's stylized behavior to make the commitment, presenting a positive serial correlation in target achievement between years: a greater likelihood of target achievement when a previous target is achieved (Indjejikian and Nanda 2002; Indjejikian and Matějka 2006; Aranda et al. 2014; Bol and Lill 2015). The serial correlation indicates that principals are rewarding agents for their stretched effort to achieve a target, making more lenient revisions that allow another achievement more likely in the following period. This is not only an *ex post* reward the favorable variance in the previous period but also an *ex ante* incentive not to discourage their effort exertion in subsequent periods. In contrast, principals penalize poor performance with an unfavorable variance (i.e., missing a target), setting future targets at least as difficult to achieve as their predecessors.

As illustrated in the prior literature, we expect that a principal's asymmetric response to previous target achievement (i.e., serial correlation of target achievements) exists consistently across types of measures. Extending the discussion, we make a further claim about the differential information asymmetry in performance measures as a moderating factor on the principal's choice over target achievability. In particular, we predict that if a principal values the information about the amount of extra effort (i.e., the size of a positive variance), as opposed to its presence *per se* (i.e., whether a target having been achieved), a marginal unit of a positive performance variance contributes to the increase of the likelihood of target achievement in a subsequent period. Otherwise, the size (or a marginal unit) of a performance variance hardly accounts for the likelihood of target achievement. This leads to the following pair of hypotheses.

H1a: *For common measures, the likelihood of target achievement in the subsequent performance period is positively associated with the size of a favorable variance.*

H1b: *For unique measures, the likelihood of target achievement in the subsequent performance period is not associated with the size of a favorable variance.*

2.2 Revision of Target Difficulty

We also examine the extent to which performance variances in a period affect the revision of the difficulty of a target in the following period.[3] For the second pair of hypotheses,

[3] To capture the target difficulty revision, we difference the performance relative to targets (PRT) in a measure between consecutive years, or $PRT_t–PRT_{t-1}$. A PRT is a plausible *ex post* difficulty measure that represents how difficult a target actually *was* (i.e., a large PRT for a higher degree of target easiness). Thus, an increase (decrease) of a PRT from a year to another suggests that a target becomes easier (more difficult). The use of the scalar measure, as opposed to a dichotomous indicator as in H1, brings greater scrutiny to the measurement of the effects on the principal's behavior. Conceptually, the approach is similar to that of Mahlendorf et al.'s (2014) "real target revision." Contrasting it with the change in targets ($Target_t–Target_{t-1}$, "nominal target revision" in their term), they find that achievement of a target in a previous period leads to an upward adjustment in a nominal target (i.e., the target is increased) but it does not induce a more challenging target (i.e., the difficulty remains the same). Unlike their measure of the change in perceived target difficulty which is collected from survey responses, we use performance variances as an objective measure.

still at the core of our discussion is a principal's asymmetric response to the previous performance. Also, we continue to argue that a marginal unit of a favorable variance in common measures corresponds better to an additional unit of an agent's effort than that in unique measures. In particular, honoring each marginal unit of favorable variances in common measures, a principal opts to commit to partial ratcheting or partial reversal of a past variance. Inversely, the principal finds the past performance information in unique measures less informative than that in common ones and attempts to reverse larger (or even the entire) part of the previous favorable variance in the former than in the latter. This results in a smaller upward target difficulty revision for positive variances in common measures than in unique measures. On the other hand, an unfavorable variance in both types is little informative of an agent's effort. Thus, negative performance variances would incur no downward target difficulty revision for both types of measures.

H2a: *The degree of upward adjustment in target difficulty for a past positive performance variance in common measures is smaller than that in unique measures.*

H2b: *The degree of downward adjustment in target difficulty for a past negative performance variance is not different between common and unique measures.*

3. Research Setting

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 USD 100 billion, and its employees are over 200,000 worldwide.

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, budgets and management-by-objectives

(MBO) constitute important components of KC's management control system, and affect managers' incentives and behaviors to a great extent. 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.

3.1 Budgeting and Target-Setting Reviews

KC's annual planning (i.e., budgeting and target-setting) process begins when business units prepare and submit their budgets and performance targets. After the initial submission, the targets are evaluated and modified through communications up and down the hierarchy.

Before the budgets and targets are finalized, KC's in-house consulting organization is actively involved in budget evaluations, supporting the KC's top management as necessary. These consultants assist the conglomerate-level management's decisions by collecting and providing information about the various businesses and industries and, as a result, reduce the information asymmetry between managers at different organization levels (i.e., KC top management and business unit managers). Top management is expected to have as nearly as good, but certainly not as complete, local knowledge as the business unit managers.

Because of the low information asymmetry, the business unit 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 satisfy these consulting experts and ultimately their supervising managers at a higher-organization level. In developing their targets, managers often base their estimates of future performance on the past performance. They well understand that targets below past performance levels are rarely approved. This is consistent with a way to set a performance standard in a reasonably objective manner, as described by Milgrom and Roberts

(1992, 233). Once finalized, the budgets and targets are rarely renegotiated, except in extraordinary situations such as a natural disaster or a severe, unforeseen recessionary shock.

The Incentive Structure

KC maintains two separate incentive programs: short-term and long-term. The short-term incentive plan (STIP) has been in place long before the beginning of our earliest sample period (i.e., 1999). A cash-based long-term performance plan (LTPP) was introduced in 2006. Earlier, stock options were granted to a limited number of high-rank executives, but the stock option plan was abolished before FY2003.

This study focuses on KC's STIP. In the STIP, KC rewards its managers based on two equally weighted categories of performance measures: summary financial measures and key performance indicators (KPIs). All business units are evaluated in terms of the summary financial measures of sales and operating profits (or EBITDA). The KPIs are selected to be informative of the performance of core strategic tasks that are unique to each business unit. The 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, brand recognition indices, and percent on-time deliveries. Importance weightings of each constituent performance measure are also negotiated as part of the annual planning process.

Insert Figure 1 about here.

3.2 Evaluation Scheme

Performance evaluations are based on both target achievement, or performance relative to target (i.e., actual/target), and performance relative to peers or, if peers do not exist, relative to prior period performance. Target achievement, for each defined performance measure is converted into the three-level evaluation ratings of Overachieved, Achieved, or Missed. Relative performance to peers or previous periods is rated as Outstanding, Par, or Below Standard.

For each performance measure, the ratings are multipled together to create an overall rating on a nine-point scale. 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,[4] 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.

4. Research Design

4.1 Data

We conduct our analysis using a set of data about business unit managers' annual performance evaluations. The dataset consists of 1,214 performance measure-years that the conglomerate used for annual performance evaluation of its twenty-nine business units (and their managers) from 2005 to 2009. From the full set of observations, we exclude six observations with missing information. Out of the remaining 1,208 performance measure-years, 239 measure-years are isolated observations; the measures are used for only one year or there is a time gap. These observations are excluded from the sample as the main analysis of this research involves serial correlations from year to year. From the observations available for at least two consecutive years, we remove 292 measure-years that are the first-time observations in a respective measure-year series. Finally, to mitigate the effects of huge outliers, we eliminate 69 observations that are

[4] S is for Superior.

in the 5% at both ends in the size of target difficulty revision: i.e., $PRT_t – PRT_{t-1}$ less than -0.6347 or greater than 0.7185. As a result, the final dataset includes 607 observations, each of which is provided with all necessary, relevant information for subsequent multivariate analyses of the inter-temporal target revision.

Insert Table 1 about here.

4.2 Measures

Performance Relative to Target (PRT) and Variance. Performance-relative-to-target (*PRT*) is computed as the ratio of actual performance relative to a target denoted in a contracted measure, written as $actual\ performance_{it}/target_{it}$. Herein, *i* represents a performance measure used in STIP, *t* denotes a year. Then, performance variance ($Variance_{it}$) is computed as $PRT_{it} – 1$.

Target Achievement and Extraordinary Performance. For performance targets of which *Variance* is greater than or equal to zero, an indicator variable $DAchieved_t$ ($DMissed_t$) is assigned one (zero) and zero (one) otherwise. In addition, excess positive or negative performance refers to as a performance variance greater than 0.2 in the absolute value (i.e., $|\text{Variances}| \geq 0.2$); or *PRT* that exceeds (falls below) 1.2 (0.8) of a target level.[5] Then, we assign one to indicator variables accordingly, $DPEx_t$ for positive excess and $DNEx_t$ for negative excess. Accordingly, the observations of which performance variances are between -0.2 and 0.2 are referred to as normal performance.

Target Difficulty Revision. *PRT* (and accordingly *Variance*) is the extent to which a manager achieves a specific target. It is an *ex post* measure for realized target difficulty. A target

[5] Observations of extraordinary performance are not extreme observations or outliers. Extraordinary performance in this paper refers to the performance outside the ±0.2 variance range where approximately 70% of our observations are populated. We see that the range generally accords with the commonsensible definition of normal performance.

with *PRT* of 120% (or *Variance* of 0.2), for example, indicates not only that actual performance exceeds the target by 20% but also that meeting the target is relatively easier to achieve than another performance target with, for example, *PRT* of 105% or 95% (or *Variance* of 0.05 or -0.05). We compare PRTs between years to understand how target difficulties are revised. Specifically, the change in the target difficulties is measured as PRT_t—PRT_{t-1}. As a result, a positive (negative) value indicates an easier (more difficult) target as compared with the previous target.

Performance Measure Types. The KC's STIP utilizes diverse performance measures of different attributes. In order to deal with this diversity, we sort performance measures into common and unique (or less common) (e.g., Lipe & Salterio, 2000) based on the relative frequency of adoption. The classification is intuitively appealing: unique measures are either specific to business units (i.e., literally "unique") or less frequently used, while common measures are those adopted more frequently than the others. In particular, we first consider performance measures that are customized and specific to business units. For example, production capacity for a certain line of product and the degree of supplier diversification are used in a single or only a few business units. The identification strategy finds 243 unique measure-years out of total 1,208. In addition, we consider relatively less frequently used measures as unique. Specifically, we compute the median frequency of adoption excluding the previously identified 243 business unit specific, unique measures: 62 times for market share measures. Then, we classify the measures adopted more or equally frequently than 62 times as common measures. They include profit (adopted 255 times; for the following measures, the figures in parentheses represent their respective adoption frequencies), sales (163), business composition (111), KC's employee survey results (132), customer satisfaction (86), and market share (62). Accordingly, the less frequently adopted measures–business initiatives (38), human resource management (35), research and development (33), brand recognition (31), and cost management (19)–are classified as unique measures. This process identifies additional 156 measure-years. In total, we identify 399 unique-measure-years and 809 common-measure-years. Table 2 summarizes the classification of performance measures.

Insert Table 2 about here.
----

5. Empirical Results

5.1 Descriptive Statistics

Table 3 describes our sample. Panel A shows that the observations are balanced across the sample period from 2005 to 2009. On average a business unit uses nine performance measures in its STIP. The panel also shows a gently rising trend of *PRT* and target achievability. While the mean (median) *PRT* increases from 0.987 (0.978) to 1.314 (1.043)[6] during the sample period, both the means and the medians are close to or slightly greater than one. For the full sample, the mean of *PRT* is 1.086 while its median (1.011) is slightly above one. The targets are generally more likely to be achieved than not; on average, 59.1% of targets are achieved.

Insert Table 3 about here.
----

[6] The mild increase of the median of *PRT* in 2009 (Δ0.021) compared with the sharp increase of the mean (Δ0.259) and the huge standard deviation (*s.d.*=1.379) suggest the likely presence of large outliers in 2009. Indeed, a greater number of huge outliers exist in 2009 compared to the other years. We attribute it to greater difficulty of target setting after an economic crisis. 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 (12.722, 11.030, and 10.930%) in the year are almost twice and three times as large as the maximum PRT (6.273) and the second largest (4.101) of all the other years respectively. Excluding the 16 outliers, the mean and the median of the other 232 observations are 1.036 and 1.023.

Panel B shows the same results broken down by type of measures. There are 809 common measure-years (67.0%) and 399 unique measure-years (33.0%). The means and the medians of PRTs are statistically not different between common measures and unique measures. However, its standard deviation in unique measures (*s.d.*=0.42) is significantly smaller than common measures (*s.d.*=0.82) at the 1% level (Δ=0.072, p=0.017). The probability of target achievement is significantly higher in unique measures than common measures (Δ=0.072, p<0.01).

Panel C shows the distribution of performance levels around the target level. At the top, we present the number of observations and the relative frequency for each performance variance interval of a one percent range from zero to positive or negative five percent. The cumulative frequencies are presented at the bottom. Also, to visualize the distributional characteristics around the target level reported in the panel, Figure 2 presents the histogram of *PRT*.

Insert Figure 2 about here.

Panel C of Table 3 and Figure 2 report abnormality in the frequency distribution of actual performance relative to a benchmark. In particular, we find a substantial divergence of the frequencies of *PRT* immediately below and above the target level for all types of measures. In the figure, the frequency distributions have a deep pit immediately below a target level (i.e., 100%) and, in contrast, the highest peak immediately above the target level. The discontinuity at a benchmark has been documented as evidence of a managers' strong incentive to achieve a target (Burgstahler and Chuk 2013; Burgstahler and Dichev 1997; Burgstahler and Eames 2006; Degeorge et al. 1999; Hayn 1995; Indjejikian, Matějka, Merchant, et al. 2014). It is of little surprise that such incentives are expected under the KC's incentive plan. In particular, KC managers can maximize the size of a bonus under KC's STIP by meeting the targets in all performance measures in the plan. Once meeting a target, they have no reason to generate large positive variances that do not contribute to the bonus in the current period but will only result in more difficult performance targets in the following period.

While the discontinuity at the benchmark performance exists consistently in all measure types, we note the degree of discontinuity is more salient in unique measures as compared to that in common measures. Meeting unique measure targets by less than a 1%-point margin (i.e., 0~1% Variance) is 8.5 times more likely than missing them by the margin (14.79% vs. 1.75%), while the difference is only 4.6 times for common measure targets (5.69% vs. 1.24%). This is consistent between 0~5% and -5~0%: 2.7 (24.56% vs. 9.02%) and 1.6 times (18.29% vs. 11.25%) for unique and common measures respectively. More importantly, we find the discrepancy in the degree of discontinuity is mainly due to the greater likelihood of unique measure targets being just-met with a small marginal performance variance than common measure targets. Notably, the probability of missing targets in unique measures (1.75%) by less than a 1%-point margin is not different from that in common measures (1.24%) (Δ=0.52%, p=0.472). On the contrary, the likelihood of *PRT* in unique measures falling in the 0~1% interval is 14.79% which is significantly higher than 5.69% for common measures (Δ=9.10%, p<0.000).

All in all, the descriptive statistics of *PRT* in Table 3 (Panels B and C) and Figure 2 provide support for our assumption about the differential characteristics between unique and common measures. We find that the mean and the median of *PRT* do not differ; its standard deviation is significantly smaller for unique measures; the likelihood of achievement is significantly greater for unique measures; and the likelihood of meeting targets by a small margin is significantly larger for unique measures. These findings are compatible with the argument that unique measures are associated with a greater degree of information asymmetry, performance manipulability, and controllable risk than common measures.

5.2 Transition Probabilities

Table 4 shows the inter-temporal relationships of performance variances between two consecutive years. The table is constructed as a stochastic matrix (also known as Markov matrix) where each cell presents the transition probability from one performance partition ($Variance_{t-1}=i$) to another in a following period ($Variance_t=j$).[7, 8]

Insert Table 4 about here.

Serial correlations of performance between consecutive years, as shown in prior literature (e.g., Aranda et al., 2014; Indjejikian and Nanda, 2002; Indjejikian et al., 2014a), are manifested as high probabilities along the shaded cells on the diagonal from the bottom left to the top right. Indeed, we see that some high probabilities are located along the diagonal. However, the linear relationship seems not evident in the table because many high probabilities (bold-faced and marked with a spade, ♠) lie off the diagonal line as well. For example, the highest probabilities in the *-20~-10%* and *40~50%* rows are 26.9% and 30.0% which are located off the diagonal. Both are in the *0~10%* column.

Column *Similar* provides the probability of the current period's variance ($Variance_t$) being similar to the previous period's variance ($Variance_{t-1}$).[9] Again, a serial correlation of

[7] It is a conditional probability denoted as $\Pr(Variance_t=j|Variance_{t-1}=i)$ where *t* represents time, and *i* and *j* represent a variance partition which each period's variance falls in. Thus, $\sum_j \Pr(Variance_t = j|Variance_{t-1} = i) = 1$.

[8] For example, the cell where the *0~10%* row ($Variance_{t-1}$) intersects with the *0~10%* column ($Variance_t$) is 43.6; this indicates that conditional on that $Variance_{t-1}$ is between 0% and 10%, the probability of $Variance_t$ ending up in the same performance partition is 43.6%.

[9] We define similar performance as $Variance_t$ being in any of three neighboring variance partitions centering the previous' period's variance partition: so to speak, $\Pr(Variance_t=j|Variance_{t-1}=i)$ for *i* and *j* such that $i-1 \le j \le i+1$.

performance would have resulted in high probabilities along the diagonal and, hence, corresponding high values across the cells in this column, which is not the case. Instead, the transition probabilities of having similar variances between two years are high only in the middle $Variance_{t-1}$ partitions from -10 to 20% (bold-faced). They are 71.9%, 72.1%, and 69.1% respectively, which are substantially higher than the conditional probabilities in the other rows in the column. The finding suggests that a serial correlation of performance exists only within the -10~20% range.

Column *Achieved* is the probability of the current target's being achieved conditional on $Variance_{t-1}$. The achievability is significantly higher for previously achieved ones (bold-faced) than for missed ones at the 1% level. This suggests an asymmetric target revision for previous performance above or below targets.

At the bottom of Table 4, we present the unconditional probabilities of performance levels (Row *Total*). This row that, regardless of the previous variance in a measure, it is most likely that performance variance sits in the 0~10% range (31.7%). It is followed by -10~0% (18.9%), 10~20% (11.3%), and -20~-10% (9.6%). In fact, the unconditional probability of performance falling within the -20~20% range sums to 71.44% overall (bold-faced). Moreover, the performance interval covers 72.68% and 69.74% of the observations with positive and negative performance variances respectively, which generally complies with our conventional wisdom criterion for normal performance. The finding forms a basis for our separation between normal versus excess performance at the ±20% cut-off points.

Overall, transition probabilities and related statistics in Table 4 show (a) that the likelihoods of target achievement in a measure between years are serially correlated, and (b) that the performance levels are also serially correlated, but they fall only within a moderate performance range.

5.3 Correlations

Table 5 presents correlations between key variables. It shows that target achievability (*$DAchieved_t$*) is positively associated with a previous period's performance (*$Variance_{t-1}$*) and an excess positive variance (*$DPEx_{t-1}$*) and negatively with previously missing targets (*$DMissed_{t-1}$*). It also shows that a current period's performance (*$Variance_t$*) is positively associated with a previous period's performance (*$Variance_{t-1}$*) and an excess positive variance (*$DPEx_{t-1}$*), while it is negatively associated with previously missing targets (*$DMissed_{t-1}$*) and an excess negative variance (*$DNEx_{t-1}$*). Following Indjejikian and Nanda (2002), we interpret the findings as the evidence of firms' asymmetric commitment to ratcheting-type target revision. Interestingly, target difficulty revision (*$Revision_t$*) is negatively associated with *$Variance_{t-1}$* and *$DPEx_{t-1}$* and positively with *$DMissed_{t-1}$* and *$DNEx_{t-1}$*. These univariate results show serial correlations of target achievements and performance variances, but they also indicate that good performance leads to more difficult targets, whereas poor performance leads to less difficult ones.

Insert Table 5 about here.

5.4 Multivariate Analysis: Target Achievability Revision

To examine the serial correlation of the likelihoods of target achievement between years, we use the following logit regression model:

$$\begin{aligned} Logodds(p = DAchieved_{i,t}) = \beta_0 &+ \beta_1 DAchieved_{i,t-1} + \beta_2 Variance_{i,t-1} \\ &+ \beta_3 DAchived_{i,t-1} \cdot Variance_{i,t-1} + \beta_4 Weight_{i,t} \\ &+ \beta_{5-8} DYear + \beta_{9-36} DBU + \varepsilon_{i,t} \end{aligned} \quad \text{(M1)}$$

Our variables of interest are target achievement (*$DAchieved_{t-1}$*), performance variance in the previous period (*$Variance_{t-1}$*), and their interaction term. We also include the importance weight of a performance measure (*Weight*[10]). All regression models include year and business unit fixed effects (*DYear* and *DBU*) to control for common macro-economic factors and business unit wide shocks. With the target achievement model (M1), we test how the type of performance measures interacts with performance levels to affect a principal's inferences regarding agents' effort exertion and subsequent target revision.

Table 6 presents the odd ratios estimated for the normal performance sample that consists of observations within ±20% performance variances. In columns, we report the results for the full normal performance sample and those for the subsamples in measures of different types (common vs. unique). In Column "All Measures," we examine the baseline relationship without separating performance measures in their types. It shows that the likelihood of target achievement is serially correlated (OR_1=1.893,[11] p=0.094) and that it is, in addition, associated with the size of a positive performance variance (OR_3=2,504.976, p=0.071). They suggest that target achievement in the last period makes the current period's target almost two times more likely achieved and an incremental 1% of positive variance in the last period increases the odds of target achievement by 25 times.

Insert Table 6 about here.
----

[10] Importance weights on performance measures contributing to bonus calculation are provided as part of dataset. They are grouped into three categories with numerical labels such that High, Medium, and Low importance weight measures carry three, two, and one respectively. The groups are split at 15% (between High and Medium) and 8% (between Medium and Low).

[11] *OR* stands for odds ratio.

With an understanding of this baseline relationship, H1 was designed to motivate a comparison of the differential effects of the sign and the size of performance variances in measures of different types. H1a expects a relationship between positive performance variances in common measures and target achievability, i.e., OR_2 (on *Variance*)=1 and OR_3 (on *DAchieved·Variance*)>1. The table provides strong evidence for the effect of performance variances as expected. In particular, the odds ratios of *DAchieved·Variance* for common measures are substantially larger than one (OR_3=693,381.4, p=0.013) while $Variance_{t-1}$ for the measures are not statistically different from one (OR_2=0.015, p=0.242). The finding suggests a huge economic significance; for example, a +5% performance variance from the target in common measures increases the estimated target achievability in the following period by approximately 16.20% (from 57.26% to 66.54%).[12] In contrast, the unique measure model does not report significance for OR_2 (OR_2=16.010, p=0.788) and OR_3 (OR_3=0.000, p=0.501). The findings support H1b.

5.4.1 Validating the Inferences from the Interaction Term

As we include an interaction term in a nonlinear model, the statistical significance of the interaction may not allow for a correct interpretation (Ai and Norton 2003; Brambor et al. 2006). To address the concern, we take three alternative approaches to understand the effects of the interaction term. First, we run alternative regression models that replace the interaction term with a quadratic term of $Variance_{t-1}$. This is a conventional approach to deal with curvilinearity (Osborne 2014). The untabulated results from the quadratic term models carry implications strongly consistent with the results in Table 6.

[12] The estimated probability of target achievement is analyzed with the "margins" command in STATA.

Second, we partition the full sample into subsamples of the combinations of performance measures and performance levels and run regressions separately.[13] In addition to the division in performance measures, we partition the observations into normal and excess performance, which results in four groups of variances: (a) Excess(-): less than -20%, (b) Normal(-):-20~0%, (c) Normal(+): 0~20%, and (d) Excess(+): greater than 20%. This is the most conservative way to test our hypotheses that involve interaction effects. In Table 7, Panel A (Panel B) compares the results between common and unique measures. In particular, the odds ratio on $Variance_{t-1}$ with normal negative variances is insignificant in both common and unique measure models (C2 and U2). However, we find that the odds ratio on $Variance_{t-1}$ is significant and greater than one in the models with normal positive variance observations in common measures (C3), but it is not in unique measures (U3). The results provide additional strong support for H1a and H1b. On the other hand, the constant odd ratios in the models for positive normal variances in unique measures (U3) are significantly larger than one (OR_0=10.939, p=0.038) while that in the common measure model is not different from one. The finding supports H1b.

Insert Table 7 about here.

Lastly, we follow Williams (2012) to examine the predicted values for interpretation of the results. Specifically, we compute the adjusted predictions at representative values (APR) and visualize them in Figure 3. Each graph refers to the estimation for a respective measure-type subsample and corresponds to each column in Table 6. In the figure, the slope of a line below zero variance represents the effect of *Variance* for the performance range (β_2) while the

[13] The business unit fixed effects are suppressed because of the small size of each subsample.

differential in the slopes, if any, would suggest the interaction effect between *Achievement* and *Variance* (β_3). Overall, Figure 3 validates the statistical findings in Tables 6 and 7.

Insert Figure 3 about here.

5.5 Multivariate Analysis: Target Difficulty Revision

In this section, we examine how performance variances affect the difficulty level of targets in the following period as compared with that of previous targets and how the relationships vary in performance measures of different types. In particular, we regress target difficulty revision ($PRT_{i,t}$–$PRT_{i,t-1}$) on performance variance in a prior year ($Variance_{i,t-1}$), indicators of performance partitions ($DMissed_{i,t-1}$, $DPEx_{i,t-1}$, and $DNEx_{i,t-1}$), an importance weight ($Weight_{i,t}$), an indicator for common measures ($DCom_i$), and interaction terms of all these variables with $Variance_{i,t-1}$ as follows:

$$\begin{aligned}
Revision_{i,t} = \beta_0 &+ \beta_1 Variance_{i,t-1} + \beta_2 DMissed_{i,t-1} \\
&+ \beta_3 DMissed_{i,t-1} \cdot Variance_{i,t-1} \\
&+ \beta_4 DPEx_{i,t-1} + \beta_5 DPEx_{i,t-1} \cdot Variance_{i,t-1} \\
&+ \beta_6 DNEx_{i,t-1} + \beta_7 DNEx_{i,t-1} \cdot Variance_{i,t-1} \\
&+ \beta_8 DCom_i \cdot Variance_{i,t-1} + \beta_9 D[C/F]_i \cdot DMissed_{i,t-1} \\
&+ \beta_{10} DCom_i \cdot DMissed_{i,t-1} \cdot Variance_{i,t-1} \\
&+ \beta_{11} DCom_i \cdot DPEx_{i,t-1} + \beta_{12} DCom_i \cdot DPEx_{i,t-1} \cdot Variance_{i,t-1} \\
&+ \beta_{13} DCom_i \cdot DNEx_{i,t-1} + \beta_{14} DCom_i \cdot DNEx_{i,t-1} \cdot Variance_{i,t-1} \\
&+ \beta_{15} DCom_i + \beta_{16} Weight_{i,t} + \beta_{17-20} DYear + \beta_{21-48} DBU + \varepsilon_{i,t}. \qquad \text{(M2)}
\end{aligned}$$

We include year and business unit dummies (*DYear* and *DBU*) to control for cross-sectional correlations across business units in each year and for business unit wide effects. The regression parameters are estimated using the Feasible Generalized Least Squares (FGLS) method to address heteroskedasticity and auto-correlations between observations.

As a negative (positive) *Revision* indicates a more difficulty (easier) new target, a negative β_1 would indicate that a positive performance variance leads to a more difficult target in the following period: i.e., “target ratcheting.” A β_1 (β_1+β_8) of minus one indicates full upward adjustment by which a new target in unique (common) measures completely reverses any previous positive variance. A zero loaded on β_1 indicates no adjustment or keeping a new target in unique measures exactly as achievable as the previous one as long as the previous target has been achieved. On the other hand, combined with the indicator for missed targets (*DMissed*), the sum of coefficients ($\beta_1+\beta_3$) represents the degree of target difficulty revision for previously missed unique measure targets. Thus, $\beta_1+\beta_3$ of minus one (zero) indicates full (no) downward target difficulty adjustment for negative variances. In the similar vein, the ratcheting for normal positive (negative) variances for common measures can be captured with $\beta_1+\beta_8$ ($\beta_1+\beta_3+\beta_8+\beta_{10}$).

In Table 8, Panel A reports the estimation for M2. Based on the estimated coefficients, Panel B tests H2a and H2b. The pair of hypotheses examines whether a performance measure type moderates a principal's target difficulty revision behavior, comparing the degree of target difficulty revision for positive and negative performance variances between common and unique measures. To test H2a (H2b) for the differential target difficulty revision for positive (negative) variances, we compare β_1 and $\beta_1+\beta_8$ ($\beta_1+\beta_3$ and $\beta_1+\beta_3+\beta_8+\beta_{10}$). Simply put, we examine the size of β_8 and $\beta_8+\beta_{10}$ for H2a and H2b respectively and the test results are presented at the bottom of Panel B.[14] A positive difference indicates a smaller upward or downward adjustment for common measures than for unique measures and a negative difference indicates the opposite. The results show that β_8 is positive and significant (β_8=0.446, p=0.005), which supports H2a—smaller upward target difficulty revision for common measures than for unique measures. For negative variances, we find $\beta_8+\beta_{10}$ is -0.457 which is not statistically significant at the 10% level

[14] The difference is to subtract the (sum of) coefficient(s) for common measures from those for unique measures. In particular, for positive variances the difference is β_8 ($=\beta_1+\beta_8-\beta_1$) and for negative variances it is $\beta_8+\beta_{10}$ ($=\beta_1+\beta_3+\beta_8+\beta_{10}-(\beta_1+\beta_3)$).

(p=0.113). The result supports H2b—little difference in a downward target difficulty revision for previously missed targets between common and unique measures.

Insert Table 8 about here.

5.5.1 Tests with Partitioned Subsamples

Finally, we run regressions separately for subsamples of different performance measures, discarding the interaction terms. Table 9 presents the results of model estimation in Panel A and tests of coefficients in Panel B. The results generally confirm the findings in Table 8, showing the clear contrasts in target revision behaviors between performance measure types. First, the results demonstrate the differential target revision behavior responding to positive variances in different measure types. β_1 for common measures is -0.420 and significantly different from minus one (p<0.000). In contrast, β_1 for unique measures (-0.977) is not different from minus one (p=0.892). The findings show that β_1 for common measures is larger than that for unique measures; that is, when a previous target is exceeded, subsequent upward adjustment of target difficulty in common measures is smaller, or less severe, than that in unique measures. Thus, H2a is supported. Further, the results suggest that the positive variances in common measures are only partly removed in a following period (i.e., incomplete target ratcheting) while those in unique measures are completely reversed (i.e., full ratcheting).

Insert Table 9 about here.

Second, target revision for normal negative variances in common measures is smaller than minus one ($\beta_1+\beta_3$=-1.225, p=0.100), while that in unique measures ($\beta_1+\beta_3$=-0.204) is statistically greater than minus one (p=0.005) and not different from zero (p=0.475, untabulated). The results suggest that downward target difficulty adjustment for negative variances in common measures is more than enough to recover fully from a previous negative variance. However, for unique measures downward adjustment is minimal, if not non-existent, which indicates a penalty. Altogether, the finding rejects H2b.

6. Conclusion

We hypothesize and test whether the type of performance measures affects ratcheting-type target-setting behaviors. To this end, we sort the types of performance measures, run two rounds of multivariate analyses for target achievement (Model 1) and target difficulty revision (Model 2), and find the following.

First, managers ratchet targets with partial upward adjustments for normal positive variances in common measures, which is in contrast with the full upward ratcheting made for unique measures. This finding suggests that the conglomerate values the information content about a manager's effort level in performance variances in common measures and chooses not to fully incorporate the information in new targets. In contrast, the firm seems to eliminate positive variances in unique measures with full ratcheting. With little knowledge about the effort-to-performance relationship in the type of measures, it may use past performance as a sole benchmark and apply a full target revision rule. Interestingly, we also find that achievement of targets in unique measures is still significantly serially correlated between years. This finding suggests that even the full ratcheting does not eliminate, if any, slacks in unique measures completely.

Second, our analysis of managers' downward revisions for normal negative variances displays contrasts between measure types. In particular, for common measures, in contrast with partial upward revision in target difficulty for normal positive variances, the firm makes full (even larger) downward revisions for normal negative variances in common measures, maintaining target difficulty at conventional levels. On the other hand, for unique measures, we find that the conglomerate requires a new target as difficult as the previous one. This can be interpreted such that the firm understands potentially large performance slack in unique measures (i.e., greater information asymmetry, manipulability, and controllable risk) and penalizes poor performance in the measures with little or incomplete downward adjustment in target difficulty.

Despite these interesting and evident findings, it must be acknowledged that our study is subject to limitations. First, our data are from practically one large conglomerate. The data are collected for 133 business-unit-year observations from 29 corporations and large business units for five years. These business units 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. Second, we cannot rule out alternative explanations that may better explain our findings including target revision behavior varying with the type of performance measures. Differential informativeness is not the only attribute of performance measures. Field-based research components such as interviews or experimental research would help to build a more solid foundation of theory about firms' target revision process and to add credibility.

References

Ai, C., and E. C. Norton. 2003. Interaction Terms in Logit and Probit Models. *Economics Letters* 80 (1): 123–129.

Aranda, C., J. Arellano, and A. Davila. 2014. Ratcheting and the Role of Relative Target Setting. *The Accounting Review* 89 (4): 1197–1226.

Bain & Company. 2015. *Management Tools & Trends 2015*.

Baker, G. 2002. Distortion and risk in optimal incentive contracts. *Journal of Human Resources*: 728–751.

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

Banker, R. D., and S. M. Datar. 1989. Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation. *Journal of Accounting Research* 27 (1): 21–39.

Baron, D. P., and D. Besanko. 1984. Regulation, asymmetric information, and auditing. *The RAND Journal of Economics*: 447–470.

Bol, J. C., and J. B. Lill. 2015. Performance Target Revisions in Incentive Contracts: Do Information and Trust Reduce Ratcheting and the Ratchet Effect? *The Accounting Review*: Forthcoming.

Bouwens, J., and P. Kroos. 2011. Target ratcheting and effort reduction. *Journal of Accounting and Economics* 51 (1): 171–185.

Bouwens, J., and L. van Lent. 2006. Performance measure properties and the effect of incentive contracts. *Journal of Management Accounting Research* 18 (1): 55–75.

Brambor, T., W. R. Clark, and M. Golder. 2006. Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis* 14 (1): 63–82.

Burgstahler, D., and E. Chuk. 2013. *What Have We Learned About Earnings Management ? Correcting Disinformation about Discontinuities*.

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

Burgstahler, D., and M. Eames. 2006. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. *Journal of Business Finance & Accounting* 33 (5–6): 633–652.

Chow, C. W., J. C. Cooper, and W. S. Wailer. 2012. Participative Budgeting : Effects Scheme on of and a Pay Information Slack and Asymmetry Performance 63 (1): 111–122.

Courty, P., and G. Marschke. 2004. An empirical investigation of gaming responses to explicit performance incentives. *Journal of Labor Economics* 22 (1): 23–56.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings Management to Exceed Thresholds. *The Journal of Business* 72 (1): 1–33.

Demski, J. S. 1998. Performance measure manipulation. *Contemporary Accounting Research* 15 (3): 261–285.

Demski, J. S., H. Frimor, and D. E. M. Sappington. 2004. Efficient manipulation in a repeated setting. *Journal of Accounting Research*: 31–49.

Dilla, W. N., and P. J. Steinbart. 2005. Relative weighting of common and unique balanced scorecard measures by knowledgeable decision makers. *Behavioral Research in Accounting* 17 (1): 43–53.

Feltham, G. A., and J. Xie. 1994. Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations. *The Accounting Review* 69 (3): 429–453.

Gibbs, M. J., K. A. Merchant, W. A. Van der Stede, and M. E. Vargus. 2009. Performance measure properties and incentive system design. *Industrial Relations: A Journal of Economy and Society* 48 (2): 237–264.

Guidry, F., A. J. Leone, and S. Rock. 1999. Earnings-based bonus plans and earnings management by business-unit managers 26: 113–142.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (2): 125–153.

Healy, P. M. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (1–3): 85–107.

Holmström, B. R. 1979. Moral Hazard and Observability. *The Bell Journal of Economics*: 74–91.

Holthausen, R. W., D. F. Larcker, and R. G. Sloan. 1995. Annual bonus schemes and the manipulation of earnings. *Journal of Accounting and Economics* 19 (1): 29–74.

Humphreys, K. A., and K. T. Trotman. 2011. The balanced scorecard: The effect of strategy information on performance evaluation judgments. *Journal of Management Accounting Research* 23 (1): 81–98.

Indjejikian, R. J., and M. Matějka. 2006. Organizational Slack in Decentralized Firms: The Role of Business Unit Controllers. *The Accounting Review* 81 (4): 849–872.

Indjejikian, R. J., M. Matějka, K. A. Merchant, and W. A. Van der Stede. 2014. Earnings targets and annual bonus incentives. *The Accounting Review* 89 (4): 1227–1258.

Indjejikian, R. J., M. Matějka, and J. D. Schloetzer. 2014. Target ratcheting and incentives: theory, evidence, and new opportunities. *Accounting Review* 89 (4): 1259–1267.

Indjejikian, R. J., and D. Nanda. 1999. Dynamic incentives and responsibility accounting. *Journal of Accounting and Economics* 27 (2): 177–201.

Indjejikian, R. J., and D. Nanda. 2002. Executive target bonuses and what they imply about performance standards. *The Accounting Review* 77 (4): 793–819.

Indjejikian, R. J., and D. Nanda. 2003. Reply to: dynamic incentives and responsibility accounting: a comment. *Journal of Accounting and Economics* 35 (3): 437–441.

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

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

———. 1996. *The balanced scorecard: translating strategy into action*. Boston, MA: Harvard Business Press.

Laffont, J. J., and J. Tirole. 1988. The dynamics of incentive contracts. *Econometrica: Journal of the Econometric Society*: 1153–1175.

Laffont, J. J., and J. Tirole. 1993. *A theory of incentives in procurement and regulation*. The MIT Press.

Leone, A. J., and S. Rock. 2002. Empirical tests of budget ratcheting and its effect on managers' discretionary accrual choices. *Journal of Accounting and Economics* 33 (1): 43–67.

Libby, T., S. E. Salterio, and A. Webb. 2004. The balanced scorecard: The effects of assurance and process accountability on managerial judgment. *The Accounting Review* 79 (4): 1075–1094.

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

Mahlendorf, M. D., M. Matějka, and U. Schäffer. 2014. Target ratcheting, incentives, and achievability of earnings targets. *Unpublished Working Paper, Arizona State University. October*.

Merchant, K. A., W. A. Van Der Stede, and W. A. Van der Stede. 2012. *Management control systems: performance measurement, evaluation and incentives*. Third Ed. Financial Times/Prentice Hall.

Milgrom, P. R., and J. Roberts. 1992. *Economics, organization and management*. Englewood Cliffs, NJ: Prentice-Hall International.

Murphy, K. J. 2000. Performance standards in incentive contracts. *Journal of Accounting and Economics* 30 (3): 245–278.

Osborne, J. W. 2014. *Best practices in logistic regression*. SAGE Publications.

Payne, J. W., J. R. Bettman, and E. J. Johnson. 1993. *The adaptive decision maker*. Cambridge University Press.

Prendergast, C. 2002. The tenuous trade-off between risk and incentives. *Journal of Political Economy* 110 (5): 1071–1102.

Raith, M. 2008. Specific knowledge and performance measurement. *The RAND Journal of Economics* 39 (4): 1059–1079.

Slovic, P., and D. Acphillamy. 1974. Dimensional Commensurability and Cue Utilization in Comparative Judgment 1. *Organizational Behavior and Human Performance* 11 (2): 172–194.

Slovic, P., and D. MacPhillamy. 1974. Dimensional commensurability and cue utilization in comparative judgment. *Organizational Behavior and Human Performance* 11 (2): 172–194.

Smith, M. J. 2002. Gaming nonfinancial performance measures. *Journal of Management Accounting Research* 14 (1): 119–133.

Weitzman, M. L. 1980. The“ ratchet principle” and performance incentives. *The Bell Journal of Economics* 11 (1): 302–308.

Williams, R. 2012. Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata Journal* 12 (2): 308–331.

Figure 1
KC's Short-Term Incentive Plan

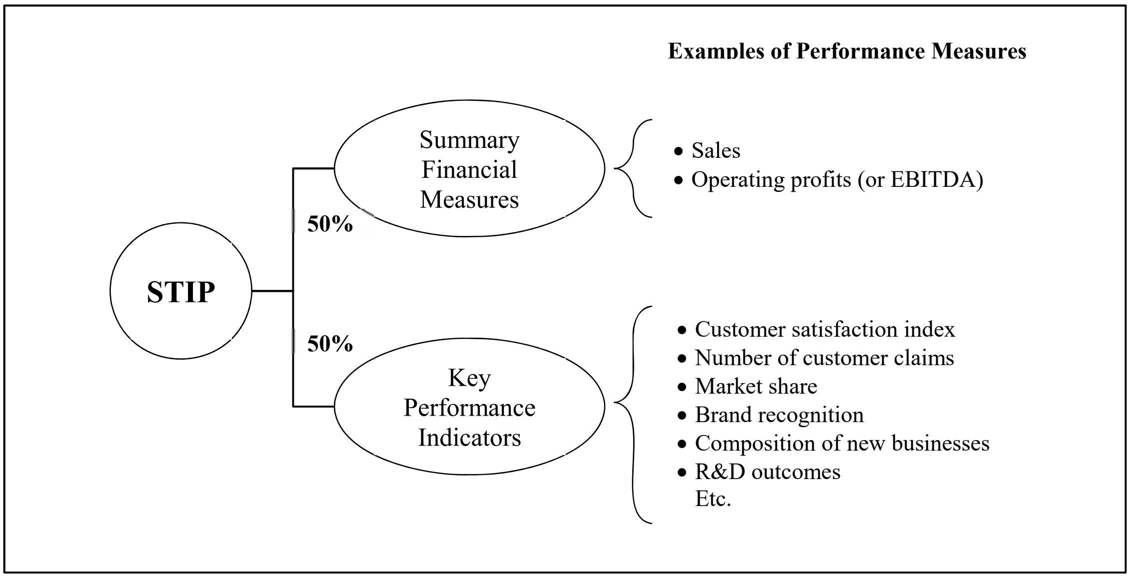
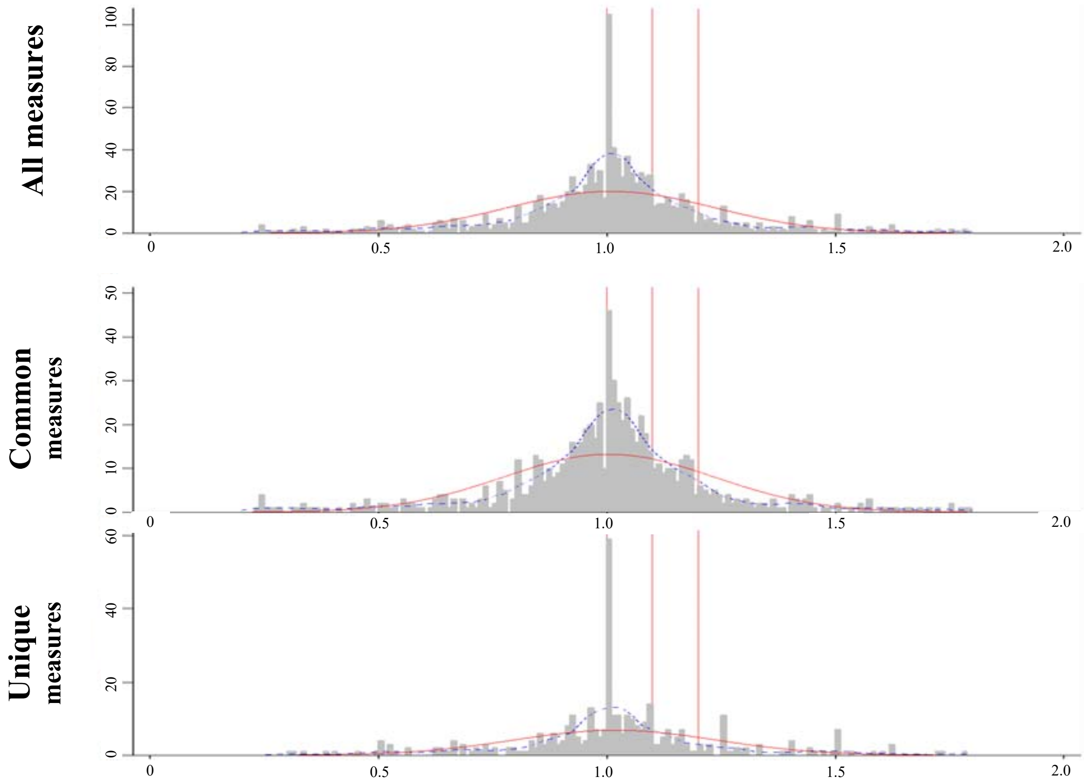


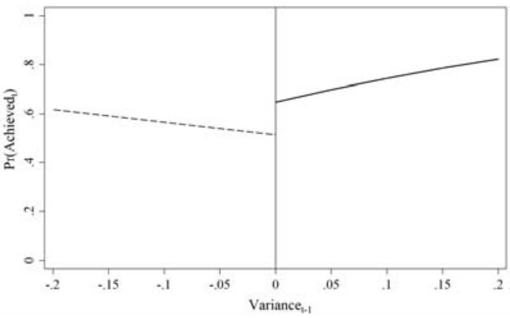
Figure 2
Distribution of Performance Relative to Target



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.

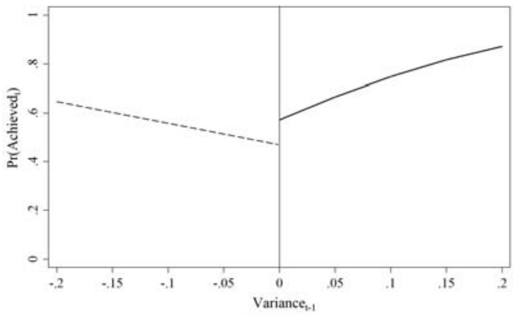
Figure 3
Effects of Previous Performance on Target Achievement

Panel A: All Measures



Panel B: Common vs. Unique Measures

Common



Unique

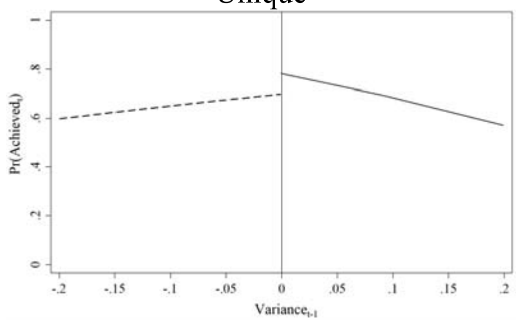


Figure 3 presents the predicted probability of target achievement for prior performance levels in performance variances. The estimation follows Williams (2012). Solid (dotted) lines are for negative (positive) variances. The differentials in the slopes between the two lines indicate the interaction effect in Model (1).

Table 1
Sample Selection for Multivariate Analyses of Serial Correlations

		Observations
Performance measure-years in the annual bonus plan		1,214
minus:	Missing information	6
Performance measure-years used for description purpose		1,208
minus:	Not used for at least two consecutive years	232
	Discontinued observations	7
	First observations	292
	Outliers in 5% at both ends [a]	69
Performance measure-years used in the following analyses		608

a. Based on the size of target revision (i.e., PRT_t—PRT_{t-1}), the cutoffs for outlying observations at the top and bottom 5% are 0.7185 and -0.6437, respectively.

Table 2
Performance Measures in the Annual Bonus Plan

		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
	Business unit specific	45	61	59	31	47	243
	Unique	**77**	**98**	**102**	**55**	**67**	**399**
Total		**229**	**249**	**253**	**229**	**248**	**1,208**

The table presents the number of business units' using respective measures in each year.
a. The median frequency of adoption (computed without 243 business unit specific measures) is 62, which constitutes the benchmark separating the common and unique performance measure types.

Table 3
Descriptive Statistics (N=1,208)

Panel A: By Year

	Year					
	2005	2006	2007	2008	2009	Total
Total no. of measures	229	249	253	229	248	1,208
Business units	27	27	27	26	29	133
Measures per BU	8.5	9.2	9.4	8.8	8.6	9.1
PRT						
Mean *PRT* [a]	0.987	1.010	1.055	1.055	1.314	1.086
Standard Deviation	0.128	0.487	0.408	0.317	1.379	0.713
Median *PRT* [a]	0.978	1.000	1.016	1.021	1.043	1.011
Achievement						
Achieved	0.459	0.546	0.609	0.629	0.706	0.591
Standard Deviation	0.499	0.499	0.489	0.484	0.457	0.492

Panel B: By Type

	N	Mean	S.D.	1Q	Med.	3Q
PRT						
Common measures	809	1.093	0.82	0.911	1.012	1.117
Unique measures	399	1.071	0.42	0.929	1.001	1.125
Difference [b]		0.023	0.40^{***}		0.011	
Achievement						
Common measures	809	0.567	0.50	0.000	1.000	1.000
Unique measures	399	0.639	0.48	0.000	1.000	1.000
Difference [b]		-0.072^{**}	0.01			

Panel C: Distribution around Zero Variance

Variance	All		Common		Unique		Δ [c]
	Number of observations (relative frequencies in %)						
-4 ~ -5%	23	(1.90)	19	(2.35)	4	(1.00)	-1.35
-3 ~ -4%	33	(2.73)	20	(2.47)	13	(3.26)	0.79
-2 ~ -3%	24	(1.99)	17	(2.10)	7	(1.75)	-0.35
-1 ~ -2%	**30**	**(2.48)**	**25**	**(3.09)**	**5**	**(1.25)**	**-1.84***
0 ~ -1%	**17**	**(1.41)**	**10**	**(1.24)**	**7**	**(1.75)**	**0.52**
0 ~ 1%	**105**	**(8.69)**	**46**	**(5.69)**	**59**	**(14.79)**	**9.10*****
1 ~ 2%	**40**	**(3.31)**	**29**	**(3.58)**	**11**	**(2.76)**	**-0.83**
2 ~ 3%	37	(3.06)	26	(3.21)	11	(2.76)	-0.46
3 ~ 4%	27	(2.24)	21	(2.60)	6	(1.50)	-1.09
4 ~ 5%	37	(3.06)	26	(3.21)	11	(2.76)	-0.46
0 ~ -20%	344	(28.48)	246	(30.41)	98	(24.56)	-5.85**
0 ~ -10%	228	(18.87)	161	(19.90)	67	(16.79)	-3.11
0 ~ -5%	**127**	**(10.51)**	**91**	**(11.25)**	**36**	**(9.02)**	**-2.23**
0 ~ 5%	**246**	**(20.36)**	**148**	**(18.29)**	**98**	**(24.56)**	**6.27****
0 ~ 10%	383	(31.71)	237	(29.30)	146	(36.59)	7.30**
0 ~ 20%	519	(42.96)	335	(41.41)	184	(46.12)	4.71

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

a. *PRT* is for *Performance Relative to Target*.

b. The equality of means, standard deviations, and medians are tested. Their differences are presented as top minus bottom.

c. Delta is the difference of the probabilities: the right (unique) minus the left (common). Column Δ in Panel exhibits the delta and the test results of the equality of the probabilities of performance variances' falling in respective ranges (i.e., Δ=0).

Table 4
Descriptive: Transition Probability

		Performance Variance @ t											[a]	[b]
	(In %)	≤ -40%	-40% ~	-30% ~	-20% ~	-10% ~	**0 ~ 10%**	~ 20%	~ 30%	~ 40%	~ 50%	> 50%	Similar	Achieved
Variance @ t-1	>50%	16.7		5.6	8.3	11.1	**19.4♠**	5.6	2.8	2.8	2.8	25.0	27.8	**58.3**
	40~50%				5.0		**30.0♠**	5.0	10.0	10.0	15.0	25.0	50.0	**95.0**
	30~40%				12.5	12.5		**37.5♠**		12.5		25.0	12.5	**75.0**
	20~30%	16.7			3.3	6.7	**36.7♠**	10.0	3.3	3.3	6.7	13.3	16.7	**73.3**
	10~20%	5.9	2.9	1.5	2.9	7.4	**36.8♠**	29.4	2.9	4.4	4.4	1.5	**69.1**	**79.4**
	0~10%	3.9	2.9	2.5	9.3	14.2	**43.6♠**	14.2	4.9	2.5		2.0	**72.1**	**67.2**
	-10~0%	3.3	2.6	3.3	9.8	28.1	**34.0♠**	6.5	5.2	2.0	2.0	3.3	**71.9**	52.9
	-20~-10%	5.1	5.1	3.8	14.1	17.9	**26.9♠**	10.3	6.4		2.6	7.7	35.9	53.8
	-30~-20%	18.5	3.7	14.8		14.8	**18.5♠**	7.4	3.7	3.7	7.4	7.4	18.5	48.1
	-40~-30%	19.0	4.8	**19.0♠**		14.3	14.3		9.5		9.5	9.5	42.9	42.9
	≤-40%	22.9	8.3	2.1	8.3	8.3	**12.5♠**	12.5	2.1		4.2	18.8	31.3	50.0
	No obs.@t-1 [c]	4.1	2.1	3.7	11.5	23.1	**30.7♠**	10.1	5.2	1.2	1.6	6.8		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		59.11

Table 4 presents the transition probability for the full sample (N=1,208). Units are per cent. ♠ indicates the largest transition probability in each row.

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 $Variance_t$ being in any of three neighboring *Variance* partitions centering the previous' period's *Variance* partition—tinted cells: so to speak, $\Pr(Variance_t = j \mid Variance_{t-1} = i)$ for i and j such that $i-1 \leq j \leq i+1$.

b. The proportion of performance measure items of which $Variance_t$ is greater than or equal to zero (i.e., $\Pr(Variance_t \geq 100 \mid Variance_{t-1})$).

c. The proportion of performance measure items of which the performance at year $t-1$ is not available (i.e., the measure is adopted first time).

Table 5
Correlations (N=608)

	Mean	Median	S.D.	1	2	3	4	5	6	7	8
1. $DAchieved_t$	0.626	1.000	0.484								
2. $Variance_t$	0.026	0.016	0.283	0.59^{***}							
3. $Revision_t$	0.031	0.011	0.236	0.38^{***}	0.51^{***}						
4. $Variance_{t-1}$	-0.006	0.000	0.257	0.31^{***}	0.62^{***}	-0.36^{***}					
5. $DMissed_{t-1}$	0.470	0.000	0.500	-0.30^{***}	-0.26^{***}	0.33^{***}	-0.58^{***}				
6. $DPEx_{t-1}$	0.100	0.000	0.301	0.20^{***}	0.38^{***}	-0.18^{***}	0.59^{***}	-0.31^{***}			
7. $DNEx_{t-1}$	0.117	0.000	0.321	-0.23^{***}	-0.29^{***}	0.30^{***}	-0.58^{***}	0.39^{***}	-0.12^{***}		
8. *DCom*	0.746	1.000	0.436	-0.09^{**}	-0.03	0.01	-0.03	0.07^{*}	-0.03	-0.01	
9. $Weight_t$	1.847	2.000	0.728	-0.09^{**}	-0.06	-0.05	-0.01	0.00	-0.01	0.01	0.51^{***}

*, **, *** indicate significance at the 10, 5, and 1 percent two-tailed confidence level respectively. The number of observations is 607.

$DAchieved_t$ is an indicator variable; 1 if the target of a performance measure is achieved in year *t* and 0 otherwise. ***$Variance_t$*** is the performance variance as compared with a target in year t: ($Actual\ Performance_t/Target_t$)–1, or PRT_t–1. ***$Revision_t$*** is the revision of target difficulties between years, computed as (PRT_t—PRT_{t-1}). ***$DMissed_{t-1}$*** 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. ***$DPEx_{t-1}$*** 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. ***$DNEx_{t-1}$*** 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. ***DCom*** is an indicator variable whose value is assigned 0 if a performance measure is used to define a business unit specific target or it is adopted less frequently than others (i.e., the number of adoption is less than its median), and 1 otherwise. ***$Weight_t$*** categorizes the importance weight of a performance measure into Low (less than 7.5%), Medium (between 8% and 12.5%,) and High (greater than 15%).

Table 6
Logit Regression Models Predicting Target Achievement in Year *t*

		Dependent: Log-odds(p=Achievement$_t$)	
OR[n]: Variables	**All Measures**	**Common**	**Unique**
[1] *DAchieved$_{t-1}$*	1.893*	1.633	1.739
	(0.722)	(0.746)	(1.446)
[2] *Variance$_{t-1}$*	0.090	0.015	16.010
	(0.273)	(0.054)	(165.173)
[3] *DAchieved$_t$*	2,504.976*	693,381.4**	0.000
· *Variance$_{t-1}$*	(10843.950)	(3.752M)	(0.002)
[4] *Weight$_t$*	0.725**	0.850	1.143
	(0.105)	(0.146)	(0.651)
[0] Constant	0.535	0.289*	3.731
	(0.320)	(0.197)	(6.957)
Observations	476	349	97
Pseudo R^2	15.44%	15.27%	18.91%

*, **, *** indicate significance at the 10, 5, and 1 percent two-tailed confidence level respectively. The table presents the odds ratios. The regressions are estimated only with the observations that are within ±20% performance variances. Clustered robust standard errors are provided in parentheses. All the models include year and business unit fixed effects. B and M stand for billion and million respectively.

DAchieved$_{t-1}$ is an indicator variable; 1 if the target of a performance measure is achieved in the previous year and 0 otherwise. ***Variance$_t$*** is the performance variance as compared with a target in year t: (*Actual Performance$_t$/Target$_t$*)–1, or *PRT$_t$*–1. ***Weight$_t$*** categorizes the importance weight of a performance measure into Low (less than 7.5%), Medium (between 8% and 12.5%,) and High (greater than 15%).

Table 7
Logit Regression Models with Performance Level Partitions

Dependent: Log-odds(p=Achievement$_t$)

	Common				Unique			
[n] Variables	**Excess(–) (C1)**	**Normal(–) (C2)**	**Normal(+) (C3)**	**Excess(+) (C4)**	**Excess(–) (U1)**	**Normal(–) (U2)**	**Normal(+) (U3)**	**Excess(+) (U4) [a]**
[1] *Variance$_{t-1}$*	49.027***	0.017	10,148.570**	9,825.239	470.485	575.822	0.154	N/A
	(68.587)	(0.054)	(38,653.31)	(58,724.83)	(2,090.025)	(,3829.410)	(0.851)	
[2] *Weight$_t$*	2.225*	1.064	0.581***	0.793	0.777	1.051	0.628	
	(0.974)	(0.234)	(0.117)	(0.725)	(0.915)	(0.542)	(0.340)	
[0] Constant	0.288	0.498	1.842	0.879	12.107	2.087	10.939**	
	(0.345)	(0.292)	(1.014)	(1.793)	(40.582)	(2.380)	(12.624)	
Observations	52	170	187	31	15	45	63	N/A
p > Chi^2	0.020	0.079	0.000	0.491	0.513	0.936	0.789	
Pseudo R^2	15.59%	5.01%	12.21%	18.12%	11.75%	2.07%	3.26%	

*, **, *** indicate significance at the 10, 5, and 1 percent two-tailed confidence level respectively. † indicates significance at the 10 percent one-tailed confidence level. The table presents the odds ratios. Clustered robust standard errors are provided in parentheses. All the models include year fixed effects. M is for million.

Variance$_t$ is the performance variance as compared with a target in year t: (*Actual Performance$_t$/Target$_t$*)–1, or *PRT$_t$*–1. ***Weight$_t$*** categorizes the importance weight of a performance measure into Low (less than 7.5%), Medium (between 8% and 12.5%,) and High (greater than 15%).

a. The regression coefficients for Models N4 and U4 cannot be estimated because of perfect prediction.

Table 8
GLS Regression Models Predicting Target Difficulty Revision with a Type Indicator

Panel A: Target Revision Including Interaction Terms with a Performance Type Indicator

		Dependent: $PRT_t - PRT_{t-1}$
β[n]	**Variables**	**Unique Vs. Common**
[1]	$Variance_{t-1}$	-0.932***
		(0.137)
[2]	$DMissed_{t-1}$	0.027
		(0.019)
[3]	$Variance_{t-1} \cdot DMissed_{t-1}$	0.221
		(0.298)
[4]	$DPEx_{t-1}$	-0.087
		(0.055)
[5]	$DPEx_{t-1} \cdot Variance$	0.428**
		(0.168)
[6]	$DNEx_{t-1}$	0.036
		(0.078)
[7]	$DNEx_{t-1} \cdot Variance$	0.594*
		(0.318)
[8]	$DCom \cdot Variance_{t-1}$	0.446***
		(0.159)
[9]	$DCom \cdot DMissed_{t-1}$	-0.017
		(0.022)
[10]	$DCom \cdot Variance_{t-1} \cdot DMissed_{t-1}$	-0.903***
		(0.345)
[11]	$DCom \cdot DPEx_{t-1}$	0.017
		(0.063)
[12]	$DCom \cdot DPEx_{t-1} \cdot Variance_{t-1}$	0.060
		(0.191)
[13]	$DCom \cdot DNEx_{t-1}$	0.015
		(0.104)
[14]	$DCom \cdot DNEx_{t-1} \cdot Variance_{t-1}$	0.086
		(0.392)
[15]	$DCom$	-0.055***
		(0.011)
[16]	$Weight_t$	-0.013***
		(0.004)
[17]	Constant	0.028
		(0.026)
	Observations	608
	Wald Chi^2	9027.57
	$p > Chi^2$	0.0000

Panel B: Test of Coefficients

Test	Unique Vs. Common
Baseline: for unique measures	
$Variance_{t-1} = -1$	-0.932
β[1]	(0.24)
$Variance_{t-1} \cdot (1 + DMissed_{t-1}) = -1$	-0.712
β[1] + β[3]	(1.37)
$Variance_{t-1} \cdot (1 + DPEx_{t-1}) = 0$	-0.505***
β[1] + β[5]	(28.36)
$Variance_{t-1} \cdot (1 + DMissed_{t-1} + DNEx_{t-1}) = 0$	-0.118
β[1] + β[3] + β[7]	(0.32)
Interaction: for common measures	
$Variance_{t-1} = -1$	-0.487***
β[1] + β[8]	(38.28)
$Variance_{t-1} \cdot (1 + DMissed_{t-1}) = -1$	-1.169
β[1] + β[3] + β[8] + β[10]	(1.40)
$Variance_{t-1} \cdot (1 + DPEx_{t-1}) = 0$	0.001
β[1] + β[5] + β[8] + β[12]	(0.00)
$Variance_{t-1} \cdot (1 + DMissed_{t-1} + DNEx_{t-1}) = 0$	-0.489***
β[1] + β[3] + β[7] + β[8] + β[10] + β[14]	(8.22)
Test of H2	
H2a: $\Delta(Variance_{t-1})$	0.446***
β[8] = 0	(7.860)
H2b: $\Delta(Variance_{t-1} \cdot (1 + DMissed_{t-1}))$	-0.457†
β[8] + β[10] = 0	(2.52)

*, **, *** indicate significance at the 10, 5, and 1 percent two-tailed confidence level respectively. † p=0.113. Panel A presents the odds ratios. The standard errors are estimated with the Feasible Generalized Least Squares (FGLS) method that addresses heteroskedasticity and auto-correlation between the panel observations. All the models include year and business unit fixed effects. Panel B presents the sum of coefficients and the respective test result as indicated. Chi^2 statistic is presented in the parenthesis.

$Variance_t$ is the performance variance as compared with a target in year t: ($Actual\ Performance_t/Target_t$)–1, or PRT_t–1. ***$DMissed_{t-1}$*** 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. ***$DPEx_{t-1}$*** 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. ***$DNEx_{t-1}$*** 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. ***DCom*** is an indicator for a common measure. ***$Weight_t$*** categorizes the importance weight of a performance measure into Low (less than 7.5%), Medium (between 8% and 12.5%,) and High (greater than 15%).

Table 9
GLS Regression Models Predicting Target Difficulty Revision with Partitioned Samples

Panel A: Target Revision with Partitioned Samples

β[n]	Variables	Common	Unique
[1]	$Variance_{t-1}$	-0.420***	-0.977***
		(0.037)	(0.166)
[2]	$DMissed_{t-1}$	-0.003	0.031
		(0.012)	(0.024)
[3]	$Variance_{t-1} \cdot DMissed_{t-1}$	-0.806***	0.774**
		(0.149)	(0.346)
[4]	$DPEx_{t-1}$	-0.083***	-0.082
		(0.032)	(0.068)
[5]	$DPEx_{t-1} \cdot Variance_{t-1}$	0.415***	0.408**
		(0.062)	(0.191)
[6]	$DNEx_{t-1}$	0.044	0.072
		(0.066)	(0.063)
[7]	$DNEx_{t-1} \cdot Variance_{t-1}$	0.692***	0.129
		(0.213)	(0.323)
[8]	$Weight_t$	-0.012***	-0.020*
		(0.003)	(0.012)
[0]	Constant	-0.019	0.011
		(0.027)	(0.118)
	Observations	452	156
	Wald Chi^2	7511.45	1326.42
	$p > Chi^2$	0.000	0.000

Panel B: Test of Coefficients

Test	Common	Unique
$Variance_{t-1} = -1$	-0.420***	-0.977
β[1]	(247.59)	(0.02)
$Variance_{t-1} \cdot (1 + DMissed_{t-1}) = -1$	-1.225*	-0.204***
β[1] + β[3]	(2.71)	(7.78)
$Variance_{t-1} \cdot (1 + DPEx_{t-1}) = 0$	-0.005	-0.569***
β[1] + β[5]	(0.01)	(36.59)
$Variance_{t-1} \cdot (1 + DMissed_{t-1} + DNEx_{t-1}) = 0$	-0.533***	-0.075
β[1] + β[3] + β[7]	(9.89)	(0.20)

*, **, *** indicate significance at the 10, 5, and 1 percent two-tailed confidence level respectively. Panel A presents the odds ratios. The standard errors are estimated with the Feasible Generalized Least Squares (FGLS) method that addresses heteroskedasticity and auto-correlation between the panel observations. All the models include year and business unit fixed effects. Panel B presents the sum of coefficients and the respective test result as indicated. Chi^2 statistic is presented in the parenthesis.

$Variance_t$ is the performance variance as compared with a target in year t: (*$Actual\ Performance_t / Target_t$*)–1, or *$PRT_t$*–1. ***$DMissed_{t-1}$*** 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. ***$DPEx_{t-1}$*** is an indicator variable; 1 if the target of a performance measure is achieved with PRT greater than 1.2 in the previous year and 0 otherwise. ***$DNEx_{t-1}$*** is an indicator variable; 1 if the target of a performance measure is underachieved with PRT less than 0.8 in the previous year and 0 otherwise. ***DCom*** is an indicator for a common measure. ***$Weight_t$*** categorizes the importance weight of a performance measure into Low (less than 7.5%), Medium (between 8% and 12.5%,) and High (greater than 15%).