Contracting on Contemporaneous versus Forward-Looking Measures: An Experimental Investigation*

ANNE M. FARRELL, *University of Illinois at Urbana-Champaign*

KATHRYN KADOUS, *Emory University*

KRISTY L. TOWRY, *Emory University*

1. Introduction

Economic theory predicts that the judicious use of nontraditional, forward-looking performance measures in incentive contracts can influence the decisions of employees whose employment horizons are shorter than the profitability horizon of the firm (Dikolli 2001).[1] Practitioner literature promotes the use of forward-looking measures as a means to facilitate employee decision making (Kaplan and Norton 1996, 2001; Case 1997; Mellor 1997; Rucci, Kirn, and Quinn 1997; Ittner and Larcker 2003; Neely and Najjar 2006), but this literature has wavered on whether further benefits can be gained from tying compensation to such measures (e.g., Kaplan and Norton 1996, 2001). In this paper, we investigate this potential dual role of forward-looking measures in incentive contracts. First, we empirically test the contracts' decision-influencing benefits as proposed by economic theory. Second and more importantly, we extend the theory by investigating whether and how forward-looking contracts facilitate the decisions of employees whose employment horizons are more closely aligned with the firm's profitability horizon.

We expect these decision-facilitating benefits to accrue because incorporating forward-looking measures into the incentive contracts of long-horizon employees more explicitly informs them of the appropriate allocation of effort across multiple periods and multiple activities (i.e., the appropriate task strategy), and thus allows them to more effectively use feedback to improve performance. Thus, employees with forward-looking contracts are able to "zero in" on the optimal strategy, leading to both a reduced need to experiment with different task strategies and greater

* Accepted by Alan Webb. This paper has benefited from comments by Alan Webb (associate editor) and anonymous reviewers, as well as from Ramji Balakrishnan, Eddy Cardinaels, Jonathan Grenier, Lynn Hannan, Lisa Koonce, Joan Luft, Liesel Mitchell, Don Moser, Mark Nelson, Mark Peecher, and Hongmei Zhao; and workshop participants at Emory University, Georgia State University, the 2005 American Accounting Association ABO Section Conference, the 2006 American Accounting Association Management Accounting Section Conference, and the 2006 Global Management Accounting Research Symposium. We thank the Research Computing staff in the Goizueta Business School at Emory University (especially Lisa Harris) for development of the program used in the experiment. We thank the Lost Dog Cafe in Arlington, Virginia and Zingerman's in Ann Arbor, Michigan for allowing us to use their creative and extensive sandwich menus in our experimental task.

Contemporary Accounting Research Vol. 25 No. 3 (Fall 2008) pp. 773–802 © CAAA
doi:10.1506/car.25.3.5

farsighted efforts. As a result of the reduced strategy experimentation, employees can direct resources they would have used for identifying the appropriate task strategy to its implementation and refinement instead. This resource redirection should result in more efficient task execution.

We conduct a multiperiod experiment in which we manipulate employees' employment horizon (long or short) and the incentive contract type (forward-looking or contemporaneous).[2] We find evidence consistent with the decision-influencing role of forward-looking contracts predicted by existing theory: use of the forward-looking contract (as compared with the contemporaneous contract) increases the farsighted efforts of short-horizon employees more than it increases those of long-horizon employees. Thus, a longer employment horizon alone motivates high levels of farsighted efforts for long-horizon employees. More notably, however, we find evidence consistent with a decision-facilitating role for long-horizon employees: those compensated with a forward-looking contract exert more farsighted effort and are more efficient in task execution. Further analysis indicates that this increase in efficiency is mediated by a reduction in experimentation across various task strategies. We obtain these results even though long-horizon employees working under the two types of contracts receive identical information about how outcomes affect future firm performance and identical performance feedback that includes both contemporaneous and forward-looking measures.

The paper's primary contribution lies in our extension of contracting theory to predict a decision-facilitating benefit of contracting on forward-looking measures. Prior theoretical literature has considered decision-influencing effects for short-horizon employees, but has not contemplated decision-facilitating effects of contracting on forward-looking measures for long-horizon employees. We also provide an empirical test of both types of benefits, and so are able to demonstrate that contracting on forward-looking measures can provide value to the firm regardless of employment horizon.

This research is also important to managers and accountants who design, implement, and evaluate performance measurement and control systems. As described earlier, practitioner literature has wavered on whether there are benefits to linking compensation to forward-looking performance measures (e.g., Kaplan and Norton 1996, 2001). Firms should be interested in our evidence that linking compensation to these measures has positive effects on employee performance beyond those gained by simply providing the measures (as in, e.g., Farrell, Kadous, and Towry 2007). Our evidence that contracting on forward-looking measures is beneficial to firms regardless of employees' horizons is particularly useful given that employment horizon is typically an employee's private information. Firms can use the insights from this research to better balance the costs of collecting and reporting forward-looking measures against the benefits derived from using the measures.

We use an experiment to address our research questions because that method allows strong causal inferences regarding the effects of different incentive contracts on performance. For example, prior archival research has examined factors that influence the contracting weights that firms place on different performance measures (e.g., Ittner and Larcker 1995, 2002; Bushman, Indjejikian, and Smith

1996; Ittner, Larcker, and Rajan 1997). However, because the contracting weights are endogenous, it is difficult to determine whether differences in firm performance are due to employee effort levels chosen in response to the weights or to the firm-level factors that led to the choice of weights (Ittner and Larcker 2001). In our study, firm characteristics are held constant across experimental conditions, allowing for stronger inferences about the effects of performance measure weights on employee performance.[3] Furthermore, while prior experimental research has examined the effects of various incentive contracts on employee performance (see Bonner and Sprinkle 2002 for a review), research in multitask settings is limited in that prior studies have generally focused on how different contracts affect the level of effort directed to one distinct action and the subsequent effect on performance.[4] We provide experimental evidence on how incentive contracts affect the allocation of limited resources across two different actions.

Finally, we introduce a novel experimental task for examining the effects of incentives on multiple, conflicting goals. The task we developed for this study is realistic and engaging. It incorporates the trade-off between the quality and quantity of production that is inherent in many real-world tasks, allowing us to clearly examine issues surrounding conflicting task demands.

The remainder of this paper is organized as follows. Section 2 explains the relevant theories and develops our hypotheses. Sections 3 and 4 describe the research design and the results of hypothesis tests, respectively. Section 5 summarizes and concludes.

2. Theory and hypothesis development

Accounting scholars have long recognized that accounting information can play both decision-influencing and decision-facilitating roles in firms (Demski and Feltham 1976; Sprinkle 2003). We rely on economics literature to make our predictions about the decision-influencing benefits of contracting on forward-looking performance measures. Economics literature, however, does not address whether and how these contracts affect the decisions of employees who intend to remain with the firm for a long time. We thus look to control theory to develop hypotheses about the decision-facilitating benefits of forward-looking contracts for long-horizon employees.

To develop our hypotheses, we use a setting in which product quality in the current period is associated with future profitability of the firm. In particular, consumers are willing to pay higher prices in the future if the firm offers high-quality products today, but prices will decline in the future if low-quality products are offered today. Thus, quality is our forward-looking performance measure. In such a setting, employees must determine how to allocate effort between two dimensions of their production task: quality and quantity. Specifically, an employee must decide whether to make more products of lower (but saleable) quality (a choice that increases contemporaneous firm financial performance at the expense of future performance) or to make fewer, higher-quality products (a choice that allows the firm to increase prices in the future, and so increases future firm financial performance at the expense of contemporaneous performance).[5]

Decision-influencing role of forward-looking measures in incentive contracts

Both academic and practitioner literature suggest that traditional financial performance measures (e.g., accounting profits) provide contemporaneous information about employee actions that affect the firm's current financial performance, but little information about actions that affect future performance. Thus, the use of contemporaneous measures in incentive contracts may motivate employees to take actions that benefit the short term to the detriment of the long term. Alternatively, other performance measures (e.g., quality) are believed to be leading indicators of future firm performance. Incorporating such forward-looking measures into incentive contracts can motivate employees to take actions that increase firm performance in the future rather than only the current period (Kaplan and Norton 1996; Ittner, Larcker, and Meyer 2003).[6] Thus, forward-looking measures play a decision-influencing role (Demski and Feltham 1976; Sprinkle 2003).

Dikolli (2001) develops an analytic model examining the decision-influencing effects of forward-looking measures. Consistent with Dikolli's 2001 model, we expect that owners can influence the effort choices of employees with short employment horizons and mitigate the goal alignment problem by incorporating forward-looking performance measures into their incentive contracts. Our first baseline hypothesis is an empirical test of this component of Dikolli's 2001 model and is illustrated in Figure 1.[7]

HYPOTHESIS 1. *Employees with short employment horizons will allocate more effort to actions that increase future firm performance when the incentive contract incorporates forward-looking performance measures than when it includes only contemporaneous measures (i.e., in Figure 1, (2) > (1)).*

Dikolli's 2001 model also proposes that the decision-influencing role of forward-looking performance measures diminishes as an employee's employment horizon approaches the profitability horizon of the firm and long-horizon employees' goals become aligned with those of the firm's owners. Thus, it is not necessary to explicitly incorporate forward-looking measures in incentive contracts for long-horizon employees; they will be motivated to take actions that increase long-term firm performance even if doing so results in lower compensation in the short term. Our second baseline hypothesis is an empirical test of this component of Dikolli's 2001 model and is illustrated in Figure 1.

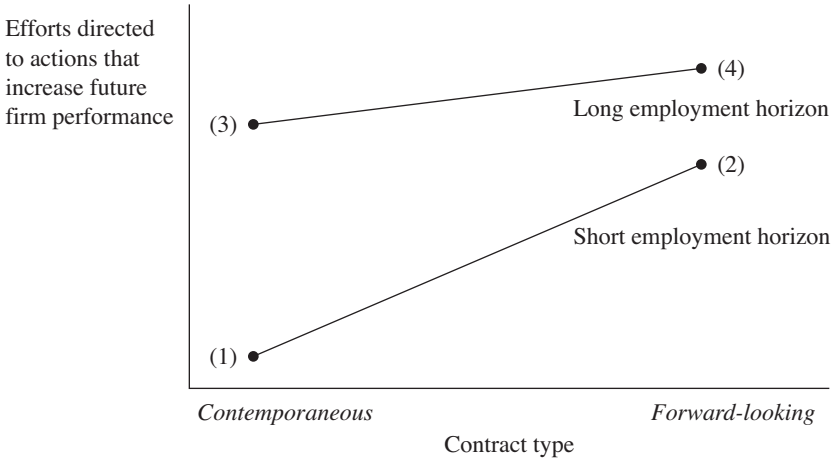
HYPOTHESIS 2. *The effect of incorporating forward-looking performance measures (as opposed to only contemporaneous measures) into incentive contracts on the allocation of efforts to actions that increase future firm performance will be greater for employees with short employment horizons than for those with long employment horizons (i.e., in Figure 1, (2) − (1) > (4) − (3)).*

Decision-facilitating role of forward-looking measures in incentive contracts

The problem that firms face with long-horizon employees is not one of goal misalignment, as it is with short-horizon employees. Rather, although long-horizon employees may be motivated to maximize long-term performance, they may have difficulty determining what current-period actions will do so. In this section, we posit a decision-facilitating effect of forward-looking contracts for long-horizon employees.

Practitioner literature promotes the use of forward-looking measures as a means to communicate to employees the strategies they need to effectively perform

Figure 1 Graph of predictions of effects of contract type and employment horizon on employees' farsighted efforts



Efforts directed to actions that increase future firm performance are measured by *average errors per sandwich* (see Table 1, note * for the variable definition). Predictions and results for these hypotheses are as follows:

Hypothesis	Prediction	Results
H1	(2) > (1)	supported
H2	(2) − (1) > (4) − (3)	supported
H4	(4) > (3)	supported

On the basis of Dikolli's 2001 model, (3) > (1). Because the effect of employment horizon across a given contract type is not our primary interest in this study, we assume Dikolli's proposition will hold in our setting, and test this assumption as a manipulation check in section 4. We do not explicitly predict differences in cells 4 and 2, but discuss actual differences in section 5.

their tasks (Kaplan and Norton 1996, 2001; Case 1997; Mellor 1997; Rucci et al. 1997; Ittner and Larcker 2003; Neely and Najjar 2006). In fact, much of this literature advocates tailoring performance measures to reflect the specific operations under the control of individual employees or divisions (Kaplan and Norton 1996, 2001; Case 1997; Rucci et al. 1997; Neely and Najjar 2006).

This literature is equivocal, however, on the extent to which linking forward-looking measures to compensation facilitates employees' decisions about what actions to take today. For example, in their 1996 book, Kaplan and Norton state, "Unless … reward and punishment are eventually tied, implicitly or explicitly, to the balanced set of objectives, measures, and targets on corporate and business scorecards, the organization will not be able to use the Balanced Scorecard as the central organizing framework for its management systems" (283). However, in their 2001 book, their position changes: "It is reasonable for companies to be cautious in tying compensation to the scorecard" (266).

Missing in this literature is the reasoning behind why simply providing forward-looking measures to employees, without tying them to compensation, may be sufficient to facilitate decision making. The underlying assumption is that when employees' goals are aligned with those of the firm's owners, and the employees have access to relevant performance measures, they will have the ability to maximize their long-term payout — an assumption consistent with economic theory (Dikolli 2001). We argue, however, that cognitive limitations will limit the employees' ability to maximize their long-term payout, and that contracting on forward-looking measures will improve their ability to do so, thus providing decision-facilitating benefits beyond those gained from only making the measures available to employees.

Consistent with economic theory, we expect long-horizon employees working under contemporaneous contracts to *attempt* to select a strategy that maximizes long-term firm performance. However, the best strategy may not be immediately obvious, and so they may need to engage in strategy experimentation. The notion of strategy experimentation is consistent with control theory, which describes the operation of self-regulating systems in psychology, engineering, economics, and other disciplines (Carver and Scheier 1982, 1998). In an organizational context, control theory suggests that employees monitor and control their performance by comparing feedback with their goals and making necessary adjustments whenever a discrepancy exists (Klein 1989). Thus, control theory suggests that long-horizon employees will use feedback to determine whether they have selected an effort allocation that maximizes long-term performance or whether a change in strategy is required.

Maximizing performance over multiple task dimensions is a difficult task. Employees have limited resources, and so increases in efforts directed to one task dimension (e.g., product quality) likely result in decreases in efforts directed to some other task dimension (e.g., product quantity). Thus, strategy experimentation can lead to erratic performance, particularly if an employee focuses first on one dimension and then on another (Carver and Scheier 1982, 1998).

Strategy selection is particularly problematic in a multiperiod setting, because maximizing long-term performance involves the analysis of the intertemporal costs

and benefits of different strategies. Decision makers are markedly poor at trading off current and future benefits (Loewenstein and Elster 1992). Prior theoretical and empirical research on multiperiod decision making finds that in the absence of a clear understanding of the optimal strategy, individuals tend to overinvest in efforts that maximize current period performance to the detriment of future performance (Herrnstein 1990; Herrnstein and Prelec 1991, 1992; Herrnstein, Loewenstein, Prelec, and Vaughan 1993; Kachelmeier and Granof 1993; Mainwaring 1997; Warry, Remington, and Sonuga-Barke 1999; Antonides and Maital 2002; Tunney and Shanks 2002) and they tend to "learn" inappropriate strategies from their task experience (Herrnstein et al. 1993).

Prior research demonstrates that specific feedback — that is, feedback that provides details about the results of a decision or a process — improves an individual's ability to choose a strategy and apply it consistently (Te'eni 1991). This research implies that specific feedback may reduce the amount of experimentation required to select the appropriate strategy, leading to less erratic behavior. In our setting, all employees have feedback in the form of both contemporaneous and forward-looking performance measures. Therefore, it is possible that long-horizon employees, regardless of how they are compensated, will effectively use feedback and engage in relatively low levels of task strategy experimentation. However, given the considerable difficulty imposed by the multiperiod setting, we expect that specific feedback will not eliminate strategy experimentation.

We extend control theory to argue that long-horizon employees will more effectively use specific feedback to maximize long-term performance when that feedback is accompanied by clear information about an appropriate task strategy. We expect that using forward-looking measures, rather than just contemporaneous measures, to determine compensation is an effective means of providing such information.[8]

We argue that incorporating forward-looking measures into incentive contracts informs employees of the appropriate allocation of effort across multiple activities to a greater extent than does providing the measures in the form of feedback. Contracts that incorporate forward-looking measures provide an explicit signal that allocating efforts to farsighted actions will maximize long-term performance, thus directing employees to the components of feedback that are relevant for performance improvements. In contrast, incentive contracts that include only contemporaneous measures provide a signal that efforts should be directed to actions that improve contemporaneous performance, which conflicts with long-horizon employees' (and firms') goals. This makes it more difficult for employees to determine what components of feedback are relevant. Furthermore, providing forward-looking measures without linking them to compensation provides only implicit signals about the appropriate effort allocation. Employees are likely to overlook these in favor of explicit signals provided in their incentive contracts or to experiment with various strategies in order to determine the one that best satisfies multiple goals.

Because forward-looking contracts make the performance-maximizing task strategy clearer, we expect that long-horizon employees compensated with such contracts will be better able to identify the appropriate strategy, and so will engage

in less strategy experimentation than will long-horizon employees with contemporaneous contracts (and, as a result, there will be less variability in outcomes). In addition, assuming that the appropriate strategy includes allocating significant effort to farsighted actions, long-horizon employees compensated with forward-looking contracts will engage in greater farsighted efforts than those compensated with contemporaneous contracts because of the explicit, goal-congruent signal about effort allocation provided in the contract. This leads to our first two hypotheses about the decision-facilitating benefits of forward-looking contracts.

HYPOTHESIS 3. *For employees with long employment horizons, those whose incentive contracts incorporate forward-looking performance measures will engage in less experimentation with task strategies than will those whose contracts include only contemporaneous measures.*

HYPOTHESIS 4. *For employees with long employment horizons, those whose incentive contracts incorporate forward-looking performance measures will allocate more effort to actions that increase future firm performance than will those whose contracts include only contemporaneous measures (i.e., in Figure 1, (4) > (3)).*

If long-horizon employees compensated with forward-looking contracts experiment less with task strategies than do those compensated with contemporaneous contracts, then these employees will be able to devote the resources that would have been used for strategy selection to strategy implementation and refinement instead. We predict that this resource redirection will manifest in increased efficiency in task execution. Thus, forward-looking contracts facilitate decision making both by shifting employee efforts in the appropriate directions (as predicted in Hypothesis 4) and by enabling employees to increase the strength of those efforts via the redirection of limited resources. This leads to our third decision-facilitating prediction.

HYPOTHESIS 5. *For employees with long employment horizons, those whose incentive contracts incorporate forward-looking performance measures will be more efficient in the execution of the task than will those whose contracts include only contemporaneous measures.*

3. Research design

Participants and experimental design

Eighty undergraduate students enrolled in upper-level accounting courses at a large university participated in the experiment. Seventy-four percent of the participants had taken three or more economics courses, and 81 percent had taken three or more accounting courses. Males comprised 57 percent of the participant pool. Participants were paid based on their performance, as described below. On average, participants earned $25.55. Generally, prior studies examining the effects of forward-looking measures on decisions have used tasks in which participants acted as

senior managers, and so they have used as participants MBA students (e.g., Lipe and Salterio 2000, 2002; Banker, Chang, and Pizzini 2004, 2006; Kaplan and Wisner 2004; Libby, Salterio, and Webb 2004; Roberts, Albright, and Hibbets 2004; Petersen and Samuels 2006; Kelly 2007) or professionals (e.g., Haywood and Stuart 2006). In our study, the use of undergraduates is appropriate because, first, our participants acted as employees rather than managers, and, second, we do not know of factors differentiating undergraduates from MBA students or professionals that would interact with our theory (see Peecher and Solomon 2001; Libby, Bloomfield, and Nelson 2002; and Liyanarachchi and Milne 2005 on the use of student participants).

Participants acted as employees making sandwiches to order for a virtual sandwich shop. We manipulated *employment horizon* (*long* or *short*) and incentive *contract type* (*forward-looking* or *contemporaneous*) between participants. There were 12 four-minute *work periods* in the task. Thus, the full design was 2 × 2 (between-participants) × 12 (within-participants). All variables are described below and the manipulations of the between-participants variables (the focus of our study) are summarized in Exhibit 1.

For the *employment horizon* independent variable, we informed each participant in the *long employment horizon* condition that he or she worked as a sandwich maker for the same sandwich shop through all work periods, and each participant in the *short employment horizon* condition that he or she worked for a different sandwich shop in each work period. All participants worked for 12 four-minute periods, but, in order to prevent end-of-task gaming, we never told participants the number of work periods.

For the *contract type* independent variable, each participant in the *contemporaneous contract* condition was paid 5 percent of the sandwich shop's revenue generated from the sale of all sandwiches that he or she produced in each work period (a contemporaneous performance measure).[9] Each participant in the *forward-looking contract* condition was paid 5 percent of the sandwich shop's revenue generated only from the sale of perfect-quality sandwiches that he or she produced in each work period (where a "perfect-quality" sandwich had no mistakes as compared to a customer's order). Note that although the *forward-looking contract* participants received pay based on a single measure (revenue from perfect sandwiches), this single measure comprises both contemporaneous (quantity) and forward-looking (quality) components. All participants were informed that (a) sandwich shop revenue for each period was computed as the number of saleable sandwiches made (where "saleable" was defined as a sandwich with fewer than four mistakes as compared to a customer's order) times that period's selling price per sandwich; (b) sandwiches with four or more mistakes had to be discarded and thus did not generate any revenue for the shop, but that such sandwiches were not included in the computation of average mistakes per sandwich; and (c) the selling price per sandwich in the first period was $5.

We informed participants in the *long employment horizon* condition that the per-sandwich price in subsequent work periods was dependent on the average number of mistakes per sandwich in preceding periods; thus, output quality is our

Exhibit 1 Experimental materials*

Now that you have experience with the sandwich-making job and the computerized ordering system, you will work as a sandwich maker for several work periods.

Long employment horizon: You have signed a contractual agreement to work with the same sandwich shop for **all of these work periods.** Your pay each period is based on the revenue the sandwich shop earns.

Short employment horizon: You have signed a series of contractual agreements, and you will **work with a different and unrelated sandwich shop for each of these work periods.** In other words, you will work with a given sandwich shop for one and only one work period, after which you will work with a different shop. You will never work with the same sandwich shop twice, and all contractual agreements have been signed and cannot be changed or cancelled for any reason. Your pay each period is based on the revenue the sandwich shop earns.

Contemporaneous contract: Specifically, your pay will equal 5% of the total revenue you generate for the sandwich shop in each work period. **You will receive this pay in real cash at the end of today's session!**

Forward-looking contract: Specifically, your pay will equal 5% of the total revenue that the shop earns from perfect sandwiches (i.e., sandwiches with exactly zero (0) mistakes) that you produce. **You will receive this pay in real cash at the end of today's session!**

Long employment horizon: The sandwich shop's revenue for a given period is computed as the number of saleable sandwiches you make times the selling price for that period. The selling price at this sandwich shop is $5.00 in the first work period. In each subsequent period, the price at this shop depends on the average number of mistakes for all sandwiches produced in the immediately preceding period, because this affects customer demand for the sandwich shop's sandwiches. Specifically, if in a given work period there is an average of:

- exactly zero (0) mistakes per sandwich, the sandwich price in the **next** period will be *10% higher*;
- greater than zero (0) but less than two (2) mistakes per sandwich, the sandwich price in the **next** period will *remain the same*;
- two (2) or more mistakes per sandwich, the sandwich price in the **next** period will be *10% lower.*

Before the start of each work period, you will be shown the new per-sandwich price for the sandwich shop.

Short employment horizon: A sandwich shop's revenue for a given period is computed as the number of saleable sandwiches you make times the selling price for that period. The selling price at your first sandwich shop is $5.00. The selling price at each of the subsequent sandwich shops you will work with has been determined in advance, and is based on customer demand for that sandwich shop's sandwiches. Before the start of each work period, you will be given the per-sandwich price for the sandwich shop you have contracted with in that period.

Finally, sandwich shops in this market require that any sandwich with four (4) mistakes or more when compared to a customer order be thrown away. Therefore, sandwiches with four (4) or more mistakes do not produce any revenue. The mistakes in sandwiches thrown away are not included in the computation of average number of mistakes.

Note:

* Materials for a given condition included the text for either the *long* or *short* employment horizon, and either the *contemporaneous* or *forward-looking* contract.

forward-looking performance measure. Specifically, if in a given work period the average number of mistakes was exactly zero per sandwich, the sandwich price in the next period would be 10 percent higher; if the average was greater than zero but less than two mistakes per sandwich, the sandwich price in the next period would remain the same; and if the average was two or more mistakes per sandwich, the sandwich price in the next period would be 10 percent lower. This pricing function was designed such that the optimal task strategy for maximizing future sandwich shop revenue was to allocate effort primarily to quality, making as many perfect sandwiches as possible, rather than making as many saleable sandwiches as possible. The pricing function parameters were based on an analysis of the task performance of 43 pilot test participants.[10]

In contrast, we informed each participant in the *short employment horizon* condition that the selling price in subsequent periods had been determined in advance and was based on customer demand at the sandwich shop for which he or she worked in a given period. In fact, the prices provided to each participant in this condition were those generated from the work of a participant in the *long employment horizon* condition under the same contract type. This yoked design equates prices across the *short-* and *long-horizon* conditions, ensuring that price, which is exogenously set in the *short-horizon* condition, does not create any inferential difficulties. We acknowledge that this design choice creates an information difference across *short-* and *long-horizon* conditions that could potentially affect inferences for Hypothesis 2, in which the *short-* and *long-horizon* conditions are compared. However, this choice was essential to the effectiveness of the employment horizon manipulation, for which we needed to induce a substantial goal-alignment problem for the *short-horizon* participants.[11]

Importantly, we operationalize our *forward-looking contract* by paying participants only for perfect sandwiches (rather than requiring high but not perfect quality) because doing so provides for economic equivalence of the two contract types in the *long employment horizon* condition. As described above, the pricing function ensures that long-term firm revenue is maximized when an employee makes as many perfect sandwiches as possible. Thus, in the *long employment horizon* condition, the optimal task strategy for participants working under a *contemporaneous* contract (and therefore paid a percentage of firm revenue each period) is to make as many perfect sandwiches as possible. By designing our *forward-looking* contract to require perfect sandwiches, we ensure that any difference across the two contract types for *long employment horizon* employees is due to a decision-facilitating effect rather than differential economic incentives. We further note that participants in both conditions received complete information on the pricing function before the task and feedback on all measures relevant to the maximization task at the end of each work period (discussed in detail later). These design choices allow us to directly test the extent to which decision-facilitating benefits accrue from contracting on forward-looking measures beyond those from simply providing the measures to employees.

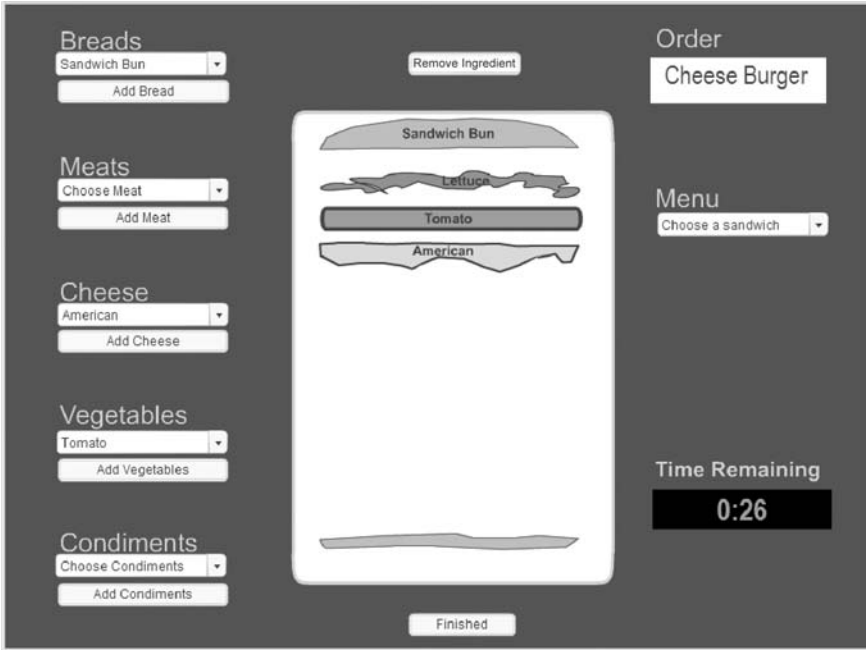
Materials and procedures

We conducted the experiment in a controlled laboratory environment using custom-designed software on stand-alone computers. Each computer ran either the *contemporaneous* or the *forward-looking contract* type condition. For both contract types, the software automatically alternated between the *long* and the *short employment horizon* conditions (i.e., the first participant using a particular computer was in the long condition, the second in short, the third in long, etc.). Prices provided to participants in the *short employment horizon* conditions were those generated by the immediately preceding participant in the *long employment horizon* condition, in accordance with the yoking procedure previously described. As participants arrived, we randomly assigned them to computers and thus to experimental conditions.

On the first screen of the computerized task, we instructed participants to assume they made their living as sandwich makers, and as such their task was to make sandwiches ordered by a sandwich shop's customers. Customer orders would be transmitted to them via a computerized ordering system. Detailed instructions about the ordering system and the sandwich-making task followed this introduction, and then participants practiced the task during a four-minute work period. For each customer order, the computer program selected from a pool of 51 preprogrammed sandwiches. To ensure that task difficulty was the same across all experimental conditions, sandwich sequences in each work period were predetermined and the same for each participant. A sample task screen is provided in Exhibit 2.

The sandwich-making task began when a customer's order appeared in the "Order" window on the computer screen. Participants used a drop-down "Menu" box to find that sandwich from among all those served at the sandwich shop, and then reviewed the sandwich's ingredient list. The ingredient list disappeared when participants moved the computer's mouse away from the "Menu" box, although they could review the ingredient list repeatedly. Participants assembled the sandwich by selecting ingredients from five drop-down menus (breads, meats, cheese, vegetables, and condiments), and as they did so, images of the selected ingredients appeared in a production space in the center of the screen. Participants could remove ingredients from the production space by using a "Remove ingredient" button. When participants completed the sandwich, they clicked on the "Finished" button. The computer program checked the assembled sandwich against its ingredient list and added it to participants' production tallies. The production space cleared, and the next customer order appeared in the "Order" window. If the last assembled sandwich had four or more mistakes, the computer program did not advance to a new sandwich in the preprogrammed sequence, but instead participants had to attempt to make the same sandwich again until the customer order was satisfied (i.e., until that sandwich was completed with fewer than four mistakes). Thus, customer perceptions of sandwich quality (and, by extension, prices) were affected by output quality, but not by production process quality.

Exhibit 2 Sample computer screen from experimental task*



Note:

* Each participant, acting as sandwich maker, completed the experimental task on screens like the one shown. A customer's sandwich order (selected from a population of 51 sandwiches) appeared in the "Order" box at the top right of the screen. The participant used the "Menu" drop-down box to find the sandwich from among all those served at the sandwich shop, and then reviewed the sandwich's ingredient list (which disappeared when the computer's mouse was moved away from the "Menu" box). As the participant assembled the sandwich using the drop-down ingredient menus on the left side of the screen, images of the ingredients appeared in the production space in the center; ingredients could be removed using the "Remove ingredient" button at the top center of the screen. When the sandwich was complete, the participant clicked on the "Finished" button at the bottom center of the screen; the sandwich was automatically checked by the computer program and added to the participant's production tally. The production space cleared, and the next customer order appeared in the "Order" box. If the last assembled sandwich had four or more mistakes, the computer program did not advance to a new sandwich in the preprogrammed sequence, but instead the participant had to attempt to make the same sandwich again until the customer order was satisfied (i.e., until that sandwich had fewer than four mistakes).

After the instructional and practice periods, participants read details of their employment relationship with sandwich shops (the *employment horizon* manipulation), and how their pay would be computed (the *contract type* manipulation). In all experimental conditions, we instructed participants that they would work as a sandwich maker for several work periods, that they would receive their pay in cash at the end of the session, and that all sandwich shops required that any sandwich with four or more mistakes be thrown away.

Participants then took a quiz to ensure they understood how prices were determined and how revenue and pay were computed. Participants could not begin the sandwich-making task until they answered all quiz questions correctly. They received reinforcement explanations of the answers even if they answered questions correctly. Successful completion of the quiz provided assurance that participants understood their employment relationship with the sandwich shops and how their pay would be computed (i.e., that our manipulations were successful).

Participants worked for 12 four-minute work periods. At the end of each period, a feedback screen displayed participant pay, sandwich shop revenue, the number of sandwiches made with zero, one, two, three, or four or more mistakes for the current period, and the per-sandwich price in the upcoming period. We provided identical feedback at the end of each work period across all experimental conditions. Thus, even participants in the *contemporaneous contract* type condition had forward-looking performance measures available. After the final work period, participants completed a post-experimental questionnaire.

4. Results

Descriptive statistics and manipulation checks

The software captured all measures used in the analysis. Descriptive statistics for dependent variables used in hypothesis tests are presented in Table 1, panel A (Hypothesis 1 and Hypothesis 2) and panel B (Hypothesis 3 to Hypothesis 5).[12] Descriptive statistics for variables used in additional analyses are presented in Table 1, panel C. Work period trends for quality and quantity efforts are depicted in Figure 2.

To check the effectiveness of our *employment horizon* manipulation, we test whether we induced a larger goal alignment problem in the *short* than in the *long employment horizon* condition by comparing efforts directed toward quality under the *contemporaneous contract*. Our measure of quality efforts is *average errors per sandwich*, computed as the simple average of the total number of errors made each period (exclusive of sandwiches thrown away) divided by the number of saleable sandwiches made; lower values for *average errors per sandwich* suggest higher levels of quality efforts.[13] Mean *average errors per sandwich* was significantly lower ($t = 5.51$, $p < 0.01$, one-tailed) in the *long employment horizon* condition (0.37) than in the *short employment horizon* condition (1.82). Thus, long-horizon participants compensated with the contemporaneous contract devoted more effort to actions that increased quality than did short-horizon participants. This suggests that our employment horizon manipulation was successful in

TABLE 1
Descriptive statistics

Panel A: Means (standard deviations) of dependent variables for Hypothesis 1 and Hypothesis 2 (decision-influencing role)

	Short employment horizon		*Long employment horizon*	
Variable	*Contemporaneous contract* (Figure 1, condition 1)	*Forward-looking contract* (Figure 1, condition 2)	*Contemporaneous contract* (Figure 1, condition 3)	*Forward-looking contract* (Figure 1, condition 4)
Average errors per sandwich[*]	1.82	0.21	0.37	0.10
	(1.01)	(0.19)	(0.60)	(0.09)
n	20	20	20	20

Panel B: Means (standard deviations) of dependent variables for Hypothesis 3 and Hypothesis 5 (decision-facilitating role)

	Long employment horizon only	
Variable	*Contemporaneous contract* (Figure 1, condition 3)	*Forward-looking contract* (Figure 1, condition 5)
Average errors per sandwich (same as panel A)[*]	0.37	0.10
	(0.60)	(0.09)
Experimentation[†]	4.25	1.85
	(3.11)	(2.11)
Range of errors per sandwich[‡]	0.81	0.39
	(0.76)	(0.26)
Efficiency[§]	758.05	1,483.71
	(724.25)	(1,485.09)
n	20	20

(The table is continued on the next page.)

TABLE 1 (Continued)

Panel C: Means (standard deviations) of variables used in additional analyses

	Short employment horizon		*Long employment horizon*	
Variable	*Contemporaneous contract* (Figure 1, condition 1)	*Forward-looking contract* (Figure 1, condition 2)	*Contemporaneous contract* (Figure 1, condition 3)	*Forward-looking contract* (Figure 1, condition 4)
Average ingredients per sandwich[#]	5.40	5.39	5.41	5.39
	(0.04)	(0.05)	(0.04)	(0.06)
Average sandwich quantity[**]	9.35	5.45	5.76	5.26
	(3.43)	(1.11)	(1.89)	(1.02)
Total discarded sandwiches[††]	4.65	1.20	0.87	0.40
	(3.36)	(1.44)	(1.42)	(0.60)
n	20	20	20	20

(The table is continued on the next page.)

TABLE 1 (Continued)

Notes:

* *Average errors per sandwich* is computed as the simple average of the total number of errors made each period (exclusive of sandwiches thrown away) divided by the number of saleable sandwiches made in that period. Lower values suggest higher levels of quality efforts.

† *Experimentation* is the response to the following item from the post-experimental questionnaire that was answered using a Likert scale of 0 (strongly disagree) to 10 (strongly agree): I changed my focus on quality versus quantity from period to period.

‡ *Range of errors per sandwich* is the difference in the maximum errors per sandwich (across all periods) and the minimum errors per sandwich (across all periods).

§ $Efficiency = \frac{Sandwich\ Quantity^2}{Total\ Errors}$

Note that efficiency is calculated using total quantity and total errors. Were we to calculate this variable for each period and then average it across periods, we would have to ignore any participant who had at least one round with zero errors.

Average ingredients per sandwich is the average number of ingredients for the actual sandwiches produced each work period.

** *Average sandwich quantity* is the simple average of the number of saleable sandwiches (i.e., exclusive of discarded sandwiches) produced per work period.

†† *Total discarded sandwiches* is the total number of sandwiches with four or more errors produced across all work periods. These sandwiches are not included in the calculation of sandwich quantity or quality.

creating a larger goal alignment problem in the *short employment horizon* condition than in the *long employment horizon* condition, and that a longer employment horizon alone can motivate high levels of farsighted efforts.

Tests of Hypotheses 1 and 2 — Decision-influencing role

Hypotheses 1 and 2 consider the decision-influencing role of forward-looking performance measures. In predicting the effort directed toward farsighted actions, this view predicts a simple main effect of *contract type* in the *short employment horizon* condition (Hypothesis 1), as well as an interaction between *contract type* and *employment horizon* (Hypothesis 2).

Table 2, panel A provides results of an analysis of variance (ANOVA) with *contract type* and *employment horizon* as the independent variables and *average errors per sandwich* as the dependent variable. Results show a significant interaction of *employment horizon* and *contract type* ($F_{1,76} = 25.14$, $p < 0.01$).[14] This suggests that the effect of the forward-looking contract on quality efforts depends on employment horizon, thus supporting Hypothesis 2. Simple effects analysis (Table 2, panel B) indicates that with a *short employment horizon*, mean *average*

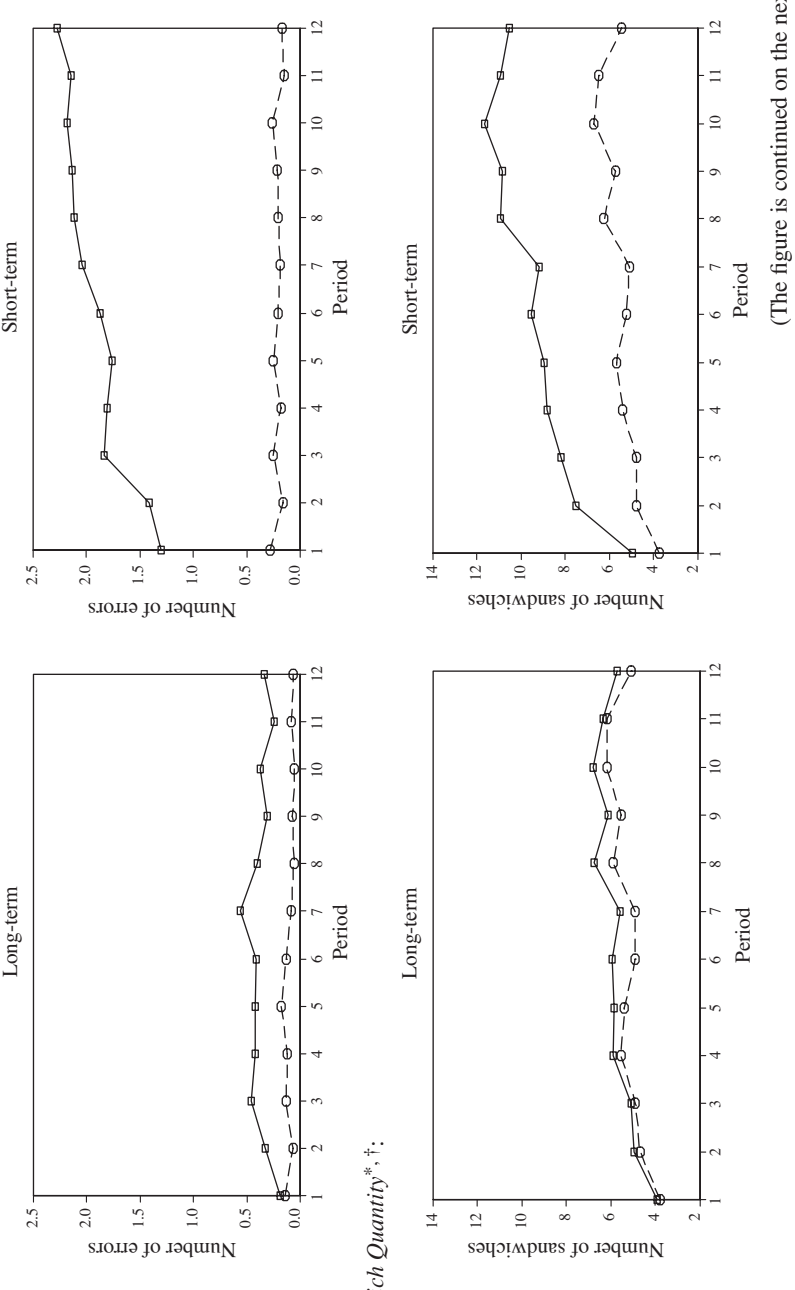


Figure 2 Work period trends for quality and quantity efforts

*Errors per Sandwich**, †:

*Sandwich Quantity**, †:

(The figure is continued on the next page.)

Figure 2 (Continued)

Notes:

* ——— *Contemporaneous* contract type condition

– – – – *Forward-looking* contract type condition

† See Table 1, notes * and ** for variable definitions.

errors per sandwich is significantly lower with the *forward-looking contract* (0.21) than with the *contemporaneous contract* (1.82; $F_{1,76} = 71.89$, $p < 0.01$), supporting Hypothesis 1.[15]

This pattern of results provides triangulating evidence consistent with the economic theory prediction that the decision-influencing benefit of incorporating forward-looking performance measures into incentive contracts depends on employees' employment horizons: the effect of compensating participants with the forward-looking contract (in contrast to the contemporaneous contract) on farsighted efforts is greater for short-horizon than for long-horizon participants.

Tests of Hypotheses 3 through 5 — Decision-facilitating role

The decision-facilitating view predicts that employees with long employment horizons will engage in less strategy experimentation when their contract incorporates forward-looking rather than only contemporaneous performance measures (Hypothesis 3), and that both farsighted efforts (Hypothesis 4) and efficiency in task execution (Hypothesis 5) will be higher for these employees.

We test Hypothesis 3 using *experimentation*, measured by participants' responses to the following item on the post-experimental questionnaire: "I changed my focus on quality versus quantity from period to period" (answered using a Likert scale ranging from 0 (strongly disagree) to 10 (strongly agree)). A planned contrast shows that *long employment horizon* participants in the *forward-looking contract* condition reported significantly less *experimentation* (1.85) than did those in the *contemporaneous contract* condition (4.25; $F_{1,76} = 5.47$, $p = 0.01$, one-tailed). Importantly, these self-reports are borne out by the observed behavior of participants. Specifically, for each *long employment horizon* participant, we calculate the *range of errors per sandwich* across all work periods (see Table 1, panel B). This variable captures the degree of strategy experimentation in that those participants who are experimenting with their strategies are likely to make many errors in some periods and few in others. Using a planned contrast, we find that mean *range of errors per sandwich* is significantly lower for *long employment horizon* participants compensated under a *forward-looking contract* (0.39) than for those under a *contemporaneous contract* (0.81; $F_{1,76} = 2.99$, $p = 0.04$, one-tailed).[16] Thus, on the basis of both participants' self-reports of behavior and their actual behavior, we find support for Hypothesis 3.

Hypothesis 4 predicts that for *long employment horizon* participants, farsighted efforts will be higher under the *forward-looking contract* than under the *contemporaneous contract*. We test Hypothesis 4 by using simple effects analysis

for our quality efforts variable (Table 2, panel B) (following from the significant interaction reported for Hypothesis 2). Results indicate that for *long employment horizon* participants, *average errors per sandwich* is lower with the *forward-looking contract* (0.10) than with the *contemporaneous contract* (0.37; $F_{1,76} = 2.08$, $p = 0.07$, one-tailed), providing support for Hypothesis 4.[17] Thus, even for long-horizon participants, the forward-looking contract results in significantly higher-quality efforts, providing evidence of a decision-facilitating role.

Hypothesis 5 predicts that for *long employment horizon* participants, the efficiency with which participants execute the sandwich-making task will be higher with the *forward-looking contract* than with the *contemporaneous contract*. Efficiency is generally defined as some measure of outputs divided by inputs. In our study, the only input is employee effort per four-minute work period. Therefore, one might consider sandwiches per period as a measure of efficiency. However,

TABLE 2
ANOVA results — *average errors per sandwich*[*]

Panel A: Results of ANOVA

Source[†]	df	Mean square	*F*	*p*[‡]
Employment horizon	1	12.14	34.20	<0.01
Contract type	1	17.71	49.88	<0.01
Employment horizon × *Contract type*	1	8.92	25.14	<0.01
Error	76	0.36		

Panel B: Simple effects of contract type on average errors per sandwich

Source[†]	df	Mean square	*F*	*p*[‡]
Long employment horizon:				
Contract type	1	0.75	2.08	0.07
Error	76	0.36		
Short employment horizon:				
Contract type	1	25.88	71.89	<0.01
Error	76	0.36		

Notes:

[*] See Table 1, note [*] for the variable definition.

[†] Because participants completed the task for multiple work periods, work period was a within-subjects variable. However, because differences across work periods are not of primary interest in this study, results are aggregated across periods and only between-subjects effects are reported in these tables.

[‡] Because hypotheses include directional predictions for contract type, *p*-values for *contract type* are reported on a one-tailed basis; all other *p*-values are two-tailed.

that measure fails to recognize that "not all sandwiches are created equal". Because effort in each work period involves a trade-off between quality and quantity efforts, it is important that any measure of efficiency consider both the quality and quantity of output.

We compare efficiency across conditions by considering the number of sandwiches produced in each condition, given the quality level — that is:

$$Efficiency = Sandwich\ Quantity / Sandwich\ Quality.$$

Because our measure of quality is *average errors per sandwich*, efficiency is operationalized as:

$$Efficiency = \frac{Sandwich\ Quantity}{\left[\frac{Total\ Errors}{Sandwich\ Quantity}\right]} \text{ or,}$$

$$Efficiency = \frac{Sandwich\ Quantity^2}{Total\ Errors}.$$

Using a planned contrast, we find that *efficiency* is significantly higher for *long employment horizon* participants compensated under a *forward-looking contract* (1,483.71) than for those under a *contemporaneous contract* (758.05; $F_{1,74} = 5.65$, $p = 0.01$, one-tailed). Thus, we find support for Hypothesis 5.[18]

Our theory suggests that the decision-facilitating benefit of increased efficiency accrues because the forward-looking contract provides an explicit signal that allocating efforts to farsighted actions will maximize long-term performance. As a result, long-horizon employees are able to "zero in" on the performance-maximizing strategy with less strategy experimentation. A benefit of the reduction in experimentation is that resources that would have been directed to strategy selection can instead be devoted to strategy implementation and refinement, therefore leading to increased efficiency. Thus, as a final step in investigating the decision-facilitating role of forward-looking contracts, we conduct a mediation analysis (Baron and Kenny 1986) to establish whether a causal chain exists from incorporating forward-looking measures into contracts to decreased strategy experimentation to increased efficiency in task execution.

First, we establish the link between *contract type* and *efficiency*. As reported previously (Hypothesis 5), *long employment horizon* participants compensated with the *forward-looking contract* are more efficient than those compensated with the *contemporaneous contract*. Second, we establish the relationship between *contract type* and *experimentation*. *Long employment horizon* participants compensated with the *forward-looking contract* reported significantly less strategy experimentation than did those compensated with the *contemporaneous contract* (Hypothesis 3). Third, we conduct an analysis of covariance (ANCOVA) with *efficiency* as the dependent variable, *contract type* as the independent variable, and

experimentation as a covariate. The effect of *experimentation* on *efficiency* is significant ($p < 0.03$, one-tailed) while the effect of *contract type* is not ($p = 0.16$, one-tailed). Thus, *experimentation* mediates the relationship between *contract type* and *efficiency*, consistent with the proposed causal chain that results in decision-facilitating benefits.

Additional analyses

In this section, we report two additional analyses to control for other variables and rule out alternative explanations for the results. The first analysis addresses the concern that levels of sandwich difficulty may have differed across conditions. All participants received the same predetermined list of sandwich orders each period. However, because not all sandwiches had the same number of ingredients, and because *average sandwich quantity* was higher in the *short employment horizon/contemporaneous contract* condition than in the other conditions (see Table 1, panel B), we rule out the possibility that quality was lower in some conditions because sandwich difficulty was higher. To do so, we compare *average ingredients per sandwich* given actual *average sandwich quantity* across conditions (Table 1, panel C). We find no differences in *average ingredients per sandwich* across conditions ($F_{3,76} = 0.56$, $p = 0.63$, two-tailed), and thus can rule out that difficulty differences caused reported quality differences.

The second analysis addresses the concern that the effect of contract type for long-horizon employees occurs because these employees behaved as if they were short-horizon employees. That is, although results reported earlier indicate that our employment horizon manipulation was successful, we also acknowledge that the long employment horizon condition is finite. Therefore, it is important to determine whether results are influenced by long-horizon participants (particularly those compensated using a contemporaneous contract) changing their behavior in anticipation of the end of the task. We conduct analyses that use the *average errors per sandwich* variable using only periods 1 through 2, 1 through 3, … , and 1 through 11. Results for each subset of the work periods are inferentially identical to those reported for all 12 work periods. Furthermore, when we partition the data into two groups — periods 1 through 6 and periods 7 through 12 — there are no inferential differences in results across the two groups, suggesting that end-of-game play does not drive reported results.

5. Conclusions

This study provides evidence that contracting on forward-looking performance measures provides benefits to firms regardless of employees' employment horizons, but that the nature of the benefits differs across horizons. For short-horizon employees, forward-looking measures play a decision-influencing role, in that incorporating these measures into incentive contracts aligns such employees' goals with those of the firm and directs employees' efforts to actions that will increase long-term firm performance. Thus, for short-horizon employees, contracting on forward-looking measures mitigates the shortsightedness that occurs when compensation is based solely on contemporaneous measures.

Because the goals of long-horizon employees are already aligned with those of the firm's owners, our results confirm the notion that incorporating forward-looking measures into incentive contracts has a smaller effect on the farsighted efforts of these employees. Importantly, though, incorporating forward-looking measures also leads to a significant increase in the farsighted efforts of long-horizon employees beyond those gained from simply providing the measures alone, consistent with a decision-facilitating role. Compensating employees on the basis of forward-looking measures simplifies long-horizon employees' multiperiod decision problem by increasing the clarity with which employees identify the performance-maximizing task strategy, thus decreasing the need for employees to experiment with various task strategies and increasing their farsighted efforts. Because resources can be redirected from strategy selection to its implementation and refinement, these employees are more efficient in task execution. Thus, while accounting researchers generally accept that incentive contracts play a decision-influencing role in firms, our results demonstrate that they can also play a decision-facilitating role, improving the performance even of those employees whose goals are more closely aligned with the firm's long-term interests. This result is particularly important in practice, because employees' goals and employment horizons are generally unobservable by firms.

One might view the efficiency gains that occur when long-horizon participants work under a forward-looking contract as a benefit arising from shifting the cost of determining task strategy from the employee to the firm. That is, to determine the contract weights for forward-looking measures, firms must undertake the costly process of developing and validating appropriate models that map forward-looking measures into long-term financial performance. In this study, we do not specify this cost, but hold it at zero across experimental conditions. However, we note that the relative efficiency with which employees and the firm can obtain this information is an important determinant of how long-horizon contracts should be constructed. For example, firms may incur these costs for strategic purposes, regardless of whether a choice is made to incorporate forward-looking measures into compensation contracts (see Ittner and Larcker 2003 for evidence on mapping performance measures in practice); furthermore, some firms may choose to determine task strategies at the firm level to avoid a duplication of efforts or undesirable variations in task strategy execution across employees. In contrast, in firms in which employees have particular expertise or access to better information, employees may be charged with determining task strategy. Therefore, future research could profitably examine the costs and benefits associated with determining task strategies at both the employee and firm levels.

Our study also provides insight into the potential benefits of employee retention plans. Our pattern of results (Table 1, panel A) suggests that when the contract includes forward-looking measures, long-horizon employees provide higher levels of farsighted efforts than short-horizon employees. Thus, if a firm could increase employees' employment horizons, firm performance gains would result even when forward-looking measures are available for contracting. Furthermore, short-horizon participants compensated with the forward-looking contract allocate levels of

effort to farsighted actions that are comparable to the levels allocated by long-horizon participants compensated with the contemporaneous contract, suggesting that the forward-looking contract can be a means to exhort short-horizon employees to act as if they had long employment horizons (consistent with economic theory). However, these results must be viewed with caution, because of informational differences across the short and long employment horizon conditions.

Although our results provide support for both the decision-influencing and decision-facilitating roles of forward-looking performance measures, our study has limitations, each of which presents an opportunity for future research. The form of the performance measure might influence results. We operationalized the forward-looking contract by paying participants a percentage of the revenue from all perfect sandwiches. Two potential issues arise from this choice. First, this contract likely makes the optimal strategy quite salient to the participants. As described earlier, this design choice has the benefit of ensuring economic equivalence across contract types for long horizon employees. Future research, however, may explore the boundaries of the decision-influencing effects we document by investigating whether less severe forward-looking contracts result in similar effects. Second, although the contract includes both contemporaneous (quantity) and forward-looking (quality) components, these are combined into one measure. Future research could address whether using separate forward-looking and contemporaneous measures would lead to even greater clarity about task strategy, as well as the more general question of whether and how the composition of performance measures in incentive contracts affects the processes leading to effective strategy development and execution.

In addition, our examination occurs in a world without uncertainty. Our firms know with certainty quality's effect on future prices and thus can determine the performance-maximizing task strategy and create an incentive system to induce employees to engage in that strategy. Employees know the length of their relationship with the firm and the basis of their compensation for the entire term of that relationship. These abstractions are important for experimental control. We encourage future research on the effects of uncertainty on the relationships we have studied.

Endnotes

1. Other remedies have been suggested for this goal incongruity problem, including integrating contract renewal bonuses and stock options into contracts (Indjejikian 1999; Murphy 1999; Jensen, Murphy, and Wruck 2004). Our study does not address the benefits and costs of other possible remedies.
2. Across all experimental conditions, we inform participants of how incentive pay will be computed before the first work period, and we provide identical performance reports that include pay and contemporaneous and forward-looking performance measures at the end of each work period. This design is consistent with Demski and Feltham's 1976 conceptualization of how a single measurement system can provide both decision-influencing and decision-facilitating benefits. Note, however, that where Demski and Feltham (1976) examine the decision-facilitating role of an accounting

measure, we more specifically examine the decision-facilitating benefits of incorporating that measure into an incentive contract.

3. Other studies examine how psychological factors affect the weights that managers place on measures when making performance evaluation judgements of employees who have already exerted effort (e.g., Lipe and Salterio 2000; Ittner et al. 2003; Banker et al. 2004; Kaplan and Wisner 2004; Libby et al. 2004; Roberts et al. 2004; Dilla and Steinbart 2005). In contrast, our study examines how weights in incentive contracts affect employees' choice of effort.
4. Exceptions include Kelly 2007 and Samuels 2005. Kelly (2007) experimentally examines the allocation of resources to capital expenditures versus research and development activities, and demonstrates that the effect of placing weight on nonfinancial measures depends on the degree to which the firm uses tangible versus intangible assets. Our study differs from Kelly's in several ways. First, we vary participants' employment horizon across conditions (an individual-level factor), whereas Kelly varies firm type (an organization-level factor). Second, our dependent measure is employee effort allocation to two dimensions of one production task, whereas Kelly's is the quality of investment decisions. Our measure captures the allocation of actual effort, whereas Kelly's captures the allocation of hypothetical resources. Finally, our design choices allow us to examine differences in task efficiency. Samuels (2005) examines how conflict between two goals (one assigned and one incentive-induced) affects goal commitment, goal conflict, effort, and performance. Our study differs from Samuels's in that we vary incentive contract type and our participants have limited resources to allocate to two actions.
5. For expositional purposes and operational ease, we assume a zero discount rate. This is a reasonable assumption in the laboratory, where all payments are made in cash at the end of the experimental session. Furthermore, this assumption has the inferential advantage of providing an unambiguous economic benchmark against which to compare results (Evans and Moser 2004).
6. Forward-looking measures can be incorporated into incentive contracts in several ways. For example, weights can be placed on two distinct measures, one that captures contemporaneous efforts and one that captures forward-looking efforts. Alternatively, a single measure can capture both contemporaneous and forward-looking efforts. Although we operationalize the latter contract in our experiment, use of the former contract would not change our hypotheses.
7. As noted previously, Dikolli's 2001 model proposes that as an employee's employment horizon approaches the profitability horizon of the firm, the goal alignment problem is reduced. Therefore, with a contract that includes only contemporaneous measures, long-horizon employees will allocate more effort to actions that increase future firm performance than will short-horizon employees (i.e., in Figure 1, (3) $>$ (1)). Because the effect of employment horizon across a given contract type is not our primary interest in this study, we assume that Dikolli's proposition will hold in our setting, and in section 4 test this assumption as a check of our employment horizon manipulation.
8. This information could be conveyed in other ways. For example, the firm could use contemporaneous contracts but communicate to employees that long-term firm performance (and thus long-term compensation) is maximized when employees take

actions to maximize quality, customer satisfaction, and/or other leading indicators of future financial performance. The degree to which employees internalize such messages likely depends on factors such as organizational trust (Kramer and Tyler 1996) or the communication mechanism used (Sinickas 2006).

9. We use revenue instead of profitability as the contemporaneous performance measure to simplify the experimental setting. This choice is reasonable if we assume that costs vary proportionally to revenue.
10. Pilot participants did not participate in the main study. To develop the pricing function, we regressed the number of sandwiches completed each period on the number of errors per sandwich in order to quantify the trade-off between quantity and quality. We used the resulting estimates to simulate the firm's revenue each period, comparing a high-quality versus high-quantity strategy. Pricing parameters were set to ensure that the high-quality strategy resulted in higher firm revenue. On the basis of the simulation, a participant using the high-quality strategy would optimally switch to a high-quantity strategy when there were only three periods remaining if the number of periods were known; however, participants were not told the number of work periods, and in the results section we report analyses to ensure that our results are not driven by participants' anticipation of the end of the experimental task.
11. If we had informed *short employment horizon* participants of the relationship between current-period quality and future-period prices at their current sandwich shop, it is possible they would have doubted the truth of our statements about the one-period contractual nature of their employment relationships. In our materials, we did not use any wording suggesting that quality was unimportant, and we included an error threshold and post-work-period feedback on both quality and quantity to all participants.
12. Because some of the data are non-normal and have heterogeneous variances, we confirm the results of all parametric hypothesis tests using rank transformations, as recommended by Conover and Iman 1981. Unless otherwise noted, results are inferentially identical with raw and rank values.
13. In our setting, output quality affects future prices. Recall that in calculating *average errors per sandwich*, our measure of output quality, we exclude discarded sandwiches. This is because discarded sandwiches are a measure of process quality. That is, when a sandwich was discarded, the participant had to remake that same sandwich again. Thus, discarded sandwiches have a negative effect on current period costs (i.e., discarded sandwiches consume valuable resources). However, they do not reach the customer, do not affect future pricing, and are therefore not an appropriate measure of future firm performance.
14. When the rank of *average errors per sandwich* is used as the dependent variable, the interaction is less significant ($F_{1,76} = 3.34$, $p = 0.07$).
15. Recall that the number of sandwiches discarded is a measure of process quality; higher numbers of discarded sandwiches are indicative of lower process quality. *Total discarded sandwiches* (Table 1, panel B) is higher for short-horizon participants in the contemporaneous contract condition (4.65) than in any other condition (means of 1.20, 0.87, and 0.40; all p-values < 0.01). The results for this process quality measure are consistent with the results for our output quality measure, *average errors per*

sandwich, as discussed above, and if we had included discarded sandwiches in the computation of *average errors per sandwich*, results would be stronger. Note, however, that our theory does not speak to the effects of forward-looking measures on process quality, a measure of contemporaneous performance.

16. Results are inferentially identical if we use the range of *sandwich quantity* instead.
17. When the rank of *average errors per sandwich* is used as the dependent variable, the simple effect of contract type in the long-horizon condition is significant at conventional levels ($F_{1,76} = 4.14$, $p = 0.02$).
18. As further evidence of efficiencies, note that long-horizon employees compensated with the forward-looking contract achieved higher quality than those compensated with the contemporaneous contract (as reported under Hypothesis 4), but they did not sacrifice volume. That is, for long-horizon participants, a planned contrast suggests that *average sandwich quantity* (Table 1, panel B) does not differ across the two contract types ($F_{1,76} = 0.06$, $p = 0.40$).

References

Antonides, G., and S. Maital. 2002. Effects of feedback and educational training on maximization in choice tasks: Experimental-game evidence. *Journal of Socio-Economics* 31 (2): 155–65.

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

Banker, R., H. Chang, and M. Pizzini. 2006. The balanced scorecard: The influence of strategy maps on decision making. Working paper, Syracuse University, Temple University, and University of Texas at Dallas.

Baron, R., and D. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6): 1173–82.

Bonner, S., and G. Sprinkle. 2002. The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society* 27 (4/5): 303–45.

Bushman, R., R. Indjejikian, and A. Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting and Economics* 21 (2): 161–93.

Carver, C., and M. Scheier. 1982. Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin* 92 (1): 111–35.

Carver, C., and M. Scheier. 1998. *On the self-regulation of behavior*. Cambridge, UK: Cambridge University Press.

Case, J. 1997. Opening the books. *Harvard Business Review* 75 (2): 118–27.

Conover, W., and R. Iman. 1981. Rank transformations as a bridge between parametric and nonparametric statistics. *American Statistician* 35 (3): 124–9.

Demski, J., and G. Feltham. 1976. *Cost determination: A conceptual approach*. Ames, IA: Iowa State University Press.

Dikolli, S. 2001. Agent employment horizons and the contracting demand for forward-looking performance measures. *Journal of Accounting Research* 39 (3): 481–94.

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

Evans, J., and D. Moser. 2004. Agency theory and incentive contracting experiments. Working paper, University of Pittsburgh.

Farrell, A., K. Kadous, and K. Towry. 2007. Does the communication of causal linkages improve effort allocations? An experimental investigation based on melioration theory. Working paper, University of Illinois at Urbana-Champaign and Emory University.

Haywood, M., and N. Stuart. 2006. Mental accounting and the balanced scorecard. Working paper, Rider University and University of South Florida.

Herrnstein, R. 1990. Rational choice theory: Necessary but not sufficient. *American Psychologist* 45 (3): 356–67.

Herrnstein, R., G. Loewenstein, D. Prelec, and W. Vaughan. 1993. Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making* 6 (3): 149–85.

Herrnstein, R., and D. Prelec. 1991. Melioration: A theory of distributed choice. *Journal of Economic Perspectives* 5 (3): 137–56.

Herrnstein, R., and D. Prelec. 1992. Melioration. In *Choice over Time*, eds. G. F. Loewenstein and J. Elster, 235–63. New York: Russell Sage Foundation.

Indjejikian, R. 1999. Performance evaluation and compensation research: An agency perspective. *Accounting Horizons* 13 (2): 147–57.

Ittner, C., and D. Larcker. 1995. Total quality management and the choice of information and reward systems. *Journal of Accounting Research* 33 (Supplement): 1–34.

Ittner, C., and D. Larcker. 2001. Assessing empirical research in managerial accounting: A value-based management perspective. *Journal of Accounting and Economics* 32 (1–3): 349–410.

Ittner, C., and D. Larcker. 2002. Determinants of performance measure choices in worker incentive plans. *Journal of Labor Economics* 20 (2): S58–S90.

Ittner, C., and D. Larcker. 2003. Coming up short on non-financial performance measurement. *Harvard Business Review* 81 (11): 88–95.

Ittner, C., D. Larcker, and M. Meyer. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review* 78 (3): 725–58.

Ittner, C., D. Larcker, and M. Rajan. 1997. The choice of performance measures in annual bonus contracts. *The Accounting Review* 72 (2): 231–55.

Jensen, M., K. Murphy, and E. Wruck. 2004. Remuneration: Where we've been, how we got to here, what are the problems, and how to fix them. Harvard Business School Negotiation, Organizations and Markets Working paper no. 04-28, Harvard University.

Kachelmeier, S., and M. Granof. 1993. Depreciation and capital investment decisions: Experimental evidence in a governmental setting. *Journal of Accounting and Public Policy* 12 (4): 291–323.

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

Kaplan, R., and D. Norton. 2001. *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Boston: Harvard Business School Press.

Kaplan, S., and P. Wisner. 2004. The judgmental effects of management communications and BSC performance dimensions on performance evaluation. Working paper, Arizona State University and the Garvin School of International Management.

Kelly, K. 2007. Feedback and incentives on nonfinancial value drivers: Effects on managerial decision making. *Contemporary Accounting Research* 24 (2): 523–56.

Klein, H. 1989. An integrated control theory model of work motivation. *Academy of Management Review* 14 (2): 150–72.

Kramer, R., and T. Tyler. 1996. *Trust in organizations: Frontiers of theory and research*. Thousand Oaks, CA: Sage Publications.

Libby, R., R. Bloomfield, and M. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27 (8): 775–810.

Libby, T., S. Salterio, and A. Webb. 2004. The balanced scorecard: The effects of assurance and process accountability on managerial judgment. *The Accounting Review* 79 (4): 1075–94.

Lipe, M., and S. Salterio. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review* 75 (3): 283–98.

Lipe, M., and S. Salterio. 2002. A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society* 27 (6): 531–40.

Liyanarachchi, G., and M. Milne. 2005. Comparing the investment decisions of accounting practitioners and students: An empirical study on the adequacy of student surrogates. *Accounting Forum* 29 (2): 121–35.

Loewenstein, G., and J. Elster, eds. 1992. *Choice over Time*. New York: Russell Sage Foundation.

Mainwaring, L. 1997. Maximisation and melioration as alternative forms of firm behavior. *Journal of Economic Behavior and Organization* 32 (3): 395–411.

Mellor, V. 1997. Communicating in an open book environment: The experience of R. R. Donnelley & Sons Company. *Strategic Communication Management* 1 (5): 22–7.

Murphy, K. 1999. Executive compensation. In *Handbook of Labor Economics*, vol. 3, eds. O. Ashenfelter and D. Card, chapter 38. Amsterdam: Elsevier Science B.V.

Neely, A., and M. Najjar. 2006. Management learning not management control: The true role of performance measurement. *California Management Review* 48 (3): 101–14.

Peecher, M., and I. Solomon. 2001. Theory and experimentation in studies of audit judgments and decisions: Avoiding common research traps. *International Journal of Auditing* 5 (3): 193–203.

Petersen, M., and J. Samuels. 2006. Missing nonstrategic targets: Are managers penalized for emphasizing important measures? Working paper, Arizona State University.

Roberts, M., T. Albright, and A. Hibbets. 2004. Debiasing balanced scorecard evaluations. *Behavioral Research in Accounting* 16: 75–85.

Rucci, A., S. Kirn, and R. Quinn. 1997. The employee-customer-profit chain at Sears. *Harvard Business Review* 76 (1): 82–97.

Samuels, J. 2005. The effect of goal conflict on effort and performance. Working paper, Arizona State University West.

Sinickas, A. 2006. Improving understanding of strategy. *Strategic Communication Management* 10 (2): 12–3.

Sprinkle, G. 2003. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society* 28 (2/3): 287–318.

Te'eni, D. 1991. Feedback as a source of control in decision support systems: An experiment with feedback specificity. *Behavior & Information Technology* 10 (5): 373–82.

Tunney, R., and D. Shanks. 2002. A re-examination of melioration and rational choice. *Journal of Behavioral Decision Making* 15 (4): 291–311.

Warry, C., B. Remington, and E. Sonuga-Barke. 1999. When more means less: Factors affecting human self-control in a local versus global choice paradigm. *Learning and Motivation* 30 (1): 53–73.