The Impact of Monitoring on Behavior and Outcome Control: Theory and Evidence on Retail Sales Productivity

Rajiv D. Banker  
Fox School of Business and Management  
Temple University  
Email: banker@temple.edu

Seok-Young Lee  
Department of Management  
Sungshin Women’s University, South Korea  
Email: sylee@sungshin.ac.kr

Jose M. Plehn-Dujowich  
Fox School of Business and Management  
Temple University  
Email: jplehn@temple.edu

October 8, 2009
The Impact of Monitoring on Behavior and Outcome Control: Theory and Evidence on Retail Sales Productivity

Abstract

We propose a principal-agent model with moral hazard to examine the impact of monitoring on behavior and outcome control. Two control schemes are considered: behavior control, which corresponds to contracting on a signal of the agent’s behavior; and behavior control in conjunction with outcome control, which corresponds to contracting on the signal and outcome of interest to the principal. The greater is the extent of monitoring chosen by the principal, the more precise is the signal. The agent exerts unobservable routine and creative effort, but since creative effort is more difficult to monitor, there is a wedge in the sensitivities of the signal and outcome to creative effort. We show that the use of outcome control and enhanced monitoring enhance the expected outcome and profits of the principal. However, if the wedge is large enough, the marginal benefit of monitoring is reduced when outcome control is used, in which case the optimal level of monitoring is lower. We test the predictions of our model by examining the effect of supervisory monitoring and the impact of implementing a monetary incentives scheme on retail sales productivity. We perform a two-stage analysis of a panel of 60 months of data from a major retailer consisting of 10 experimental stores that implemented a monetary incentives scheme and 10 control stores that did not. In the first stage, we use Data Envelopment Analysis (DEA) to compute the relative productivity of the retail stores in using their labor and capital resources to generate sales. In the second stage, we regress the logarithm of DEA scores on the level of monitoring, the provision of monetary incentives, and their interaction, yielding consistent estimators (Banker and Natarajan, 2008). In accord with our agency model, our empirical results indicate that monetary incentives and the level of supervisory monitoring have significant positive effects on retail sales productivity; and the marginal impact of monitoring and the level of monitoring are diminished when monetary incentives are provided.
I. INTRODUCTION

Organizational control theory posits that there are two forms of control, behavior and outcome (Ouchi, 1979; Eisenhardt, 1985). The optimal use of behavior versus outcome control is determined by the task’s characteristics, namely, outcome measurability and task programmability. In particular, if the task is less programmable and the outcomes are readily measured, outcome control is the appropriate strategy. Behavior and outcome control are often viewed as dichotomous; however, in practice, aspects of both forms of controls are commonly utilized (Oliver and Anderson, 1994, 1995). Furthermore, behavior control may be aided by supervisory monitoring (Eisenhardt, 1989). Yet, the theoretical literature has not considered the use of both forms of control; nor has it examined the impact of monitoring on the benefits and costs associated with each. The strong link between agency theory and organizational control theory has long been recognized (Eisenhardt, 1985, 1988, 1989). In that spirit, we develop an agency model to examine the impact of monitoring on behavior and outcome control. In doing so, we determine the conditions under which the level of monitoring serves as a complement to or a substitute for outcome control. We then test the predictions of our analytical model using a novel dataset on retail sales productivity, finding strong support.

We propose a principal-agent model with moral hazard in which a risk-neutral principal contracts with a risk-averse agent. The agent exerts routine and creative effort, both of which are unobservable. The outcome of interest to the principal depends on both types of effort. The principal monitors a noisy signal of the agent’s behavior, which is also a function of routine and creative effort. The greater is the extent of monitoring chosen by the principal, the more precise is the signal of the agent’s behavior. Monitoring is costly, at an increasing rate. The signal is relatively more accurate at measuring routine effort, reflecting the premise that creative effort is more difficult to monitor. Specifically, the sensitivity of the signal to routine effort is the same as the sensitivity of the outcome to routine effort, but the sensitivity of the signal to creative effort differs from the sensitivity of the outcome to creative effort. The wedge in the sensitivities of the signal and outcome to creative effort plays an important role in our agency theory. Two control schemes are considered: behavior control; and behavior control in conjunction with outcome control. Behavior control corresponds to contracting on the signal of the agent’s behavior. Outcome control corresponds to contracting on the outcome. Using both control schemes entails contracting on the signal in conjunction with the outcome.
Our analytical model yields four main findings. First, greater monitoring is beneficial to the principal: it enhances the expected outcome of interest to the principal and the expected profits of the principal. Behavior control entails contracting on the signal of the agent’s behavior, which is subject to monitoring. Therefore, increasing the extent of monitoring is beneficial because a rise in the precision of the signal of the agent’s behavior reduces the agency cost associated with moral hazard incurred by the principal, in accord with Banker and Datar (1989). Second, using outcome control (in conjunction with behavior control) is beneficial to the principal: it enhances the expected outcome of interest to the principal and the expected profits of the principal. Indeed, the principal cannot be made worse off by having another tool upon which to base compensation, in this case the outcome. Third, suppose the outcome is highly sensitive to creative effort. We show that the marginal benefit to the principal of monitoring is lower when the principal uses behavior control in conjunction with outcome control; similarly, the value to the principal of using outcome control when paired with behavior control diminishes as the principal increases the level of monitoring. The more sensitive is the outcome to creative effort, the greater is the wedge in the sensitivities of the signal and outcome to creative effort; thus, the less useful is the signal of the agent’s behavior as a form of control and by extension monitoring as a device by which to implement incentive compatibility. However, in the absence of outcome control, the signal of the agent’s behavior is the only means by which to elicit incentive compatibility, in which case enhanced monitoring is directly beneficial. Fourth, an implication of this last finding is that the optimal level of monitoring is lower when the principal uses outcome control (in conjunction with behavior control) if the outcome is sufficiently sensitive to creative effort.

Following in the tradition of Anderson and Oliver (1987), Oliver and Anderson (1994, 1995), and Eisenhardt (1985, 1988), we test the predictions of our analytical model by examining the effect of supervisory monitoring and the impact of the implementation of a monetary incentives scheme on retail sales productivity. The outcome of interest to the principal (i.e., the store manager or supervisor) is retail sales productivity, such that the goal of the principal is to maximize the retail sales productivity of the agent (i.e., the salesperson) net of the agent’s compensation. Outcome control corresponds to offering the salesperson monetary incentives; while behavior control corresponds to observing a noisy signal of how hard the salesperson is working, which is subject to monitoring by the supervisor.
Anderson and Oliver (1987) conceptualize sales management control as follows: outcome control encourages and rewards the results of a salesperson, such as sales volume and profitability; while behavior control encourages the inputs of a salesperson towards the selling process, such as sales call planning and customer relationship building. Prior sales management control research integrates monetary incentives into behavior-based sales management control (Oliver and Anderson, 1994). However, an important issue heretofore unexamined in this literature is how the implementation of a monetary incentives scheme influences the effectiveness of monitoring suggested by behavior control. Yet, managers’ control activities and compensation policies are essential to the motivation and direction of salespeople (Piercy et al., 2004). We examine not only the impact of monitoring and monetary incentives on retail sales productivity, but also the conflict between the two as suggested by our analytical model. In addition, little is known about the level of supervisory monitoring that interacts with monetary incentives in affecting retail sales productivity. Overall, our empirical study thereby contributes towards identifying the role of monitoring that moderates the effectiveness of behavior and outcome control so that researchers and organizations may become better informed about the use of monetary incentives (Bonner and Sprinkle, 2002).

Retailers can be broadly classified into three distinct market segments: high-end retailers like Neiman Marcus, retailers positioned in the middle such as Foley’s, and low-end retailers like Kmart (Anderson, 1999). Our research site falls in the category of a fair-priced retailer positioned in the middle of the department store industry. Price-conscious shoppers bypass them for discounters while brand-hungry consumers head to the high-end stores and other specialty stores for the top brand names. This has led to increased pressure on existing retailers such as our research site to develop new strategies to survive in this competitive environment. The strategic positioning of our research site is to command medium range prices for its merchandise by providing a reasonable level of service to its customers, going beyond simply ringing up the cash register or responding to customers’ requests. In order to achieve its strategic position, our research site introduced a monetary incentives scheme to provide services beyond customers’ expectations. This scheme was expected to lead to superior customer service and thereby improved customer satisfaction and sales performance.

We follow the two-step estimation procedure in Banker and Natarajan (2008) to study 60 months of data from 10 experimental stores (that implemented a monetary incentives scheme)
and 10 control stores (that did not implement a monetary incentives scheme). Banker and Natarajan show that the Data Envelopment Analysis (DEA) estimator of productivity obtained from the analysis of input-output data can be regressed in the second stage on contextual factors believed to contribute to productivity differences, yielding consistent estimators. We implement this technique as follows. In the first stage, we use DEA to compute the relative productivity of the retail stores in using their salespeople, inventory, and retail space to generate store sales. In the second stage, we investigate how the level of supervisory monitoring, implementation of the monetary incentives scheme, and the interaction between monitoring and monetary incentives affect the sales productivity of the 20 stores.

Our empirical results agree with the four main predictions of our analytical model, subject to the proviso that creative effort is important in customer-focused organizations (which we argue is the case at our research site). The effects on retail sales productivity of both the level of supervisory monitoring and the implementation of a monetary incentives scheme are significantly positive. Furthermore, the interaction of monitoring and monetary incentives has a significant negative effect on retail sales productivity. In other words, the marginal impact of the level of supervisory monitoring on sales productivity is lower when the monetary incentives scheme is implemented, and the value of monetary incentives decreases as the level of supervisory monitoring increases. Accordingly, we find that the level of monitoring is significantly lower when monetary incentives are provided.

The closest agency models to ours are due to Joseph and Thevaranjan (1998) and Demougin and Fluet (2001). In both Joseph and Thevaranjan and Demougin and Fluet, the principal’s only tool by which to resolve the moral hazard problem is outcome control, i.e. making compensation contingent on the outcome. By contrast, we also consider behavior control, having a signal of the agent’s behavior that is available to the principal as a contracting tool. Furthermore, in Joseph and Thevaranjan, the extent of monitoring is exogenous in the sense that it enables one type of effort to be observed perfectly, whereas it is an endogenous choice in our model that increases the precision of the agent’s signal of behavior; and the principal selects an agent on the basis of his degree of risk aversion, whereas we have no such choice. In Demougin and Fluet, both the principal and agent are risk-neutral, such that a limited liability constraint on the agent drives the moral hazard problem, whereas we follow the more traditional approach used in agency theory of having a risk-averse agent with a reservation utility; and monitoring is a
means by which to observe mistakes on the part of the agent, whereas we presume it increases the precision of the performance signal. Because neither model has behavior control, they are unable to examine the tradeoff associated with utilizing behavior versus outcome control; nor are they able to assess the role of enhanced monitoring in the context of this tradeoff. Datar et al. (2001), Holmstrom and Milgrom (1987), and Feltham and Xie (1994) examine moral hazard problems with multiple types of effort, but their frameworks do not exactly coincide with behavior and outcome control, nor do they have monitoring.

The remainder of this paper is organized as follows. Section II presents our principal-agent model of behavior and outcome control in the presence of monitoring. Section III discusses our empirical hypotheses drawing upon the literature and the predictions of our analytical model. Section IV describes the data and the estimation method utilized for the empirical analysis. Section V contains the results of our empirical analysis. Section VI summarizes and concludes. An Appendix provides the proofs of propositions.

II. ANALYTICAL MODEL

A risk-neutral principal hires a risk-averse agent. The agent exerts two types of unobservable effort, routine \( e_1 \) and creative \( e_2 \). The utility cost of exerting effort is \( C(e_1, e_2) = c_1 e_1^2 / 2 + c_2 e_2^2 / 2 \), which is linearly additive as in Feltham and Xie (1994), Joseph and Thevaranjan (1998), and Datar et al. (2001). The outcome of interest to the principal is

\[
(1) \quad x(e_1, e_2) = e_1 + \lambda e_2 + \epsilon_x,
\]

where \( \epsilon_x \) is a normal shock with mean zero and variance \( \text{Var}(\epsilon_x) = \sigma_x^2 \), and \( \lambda > 0 \) is the sensitivity of the outcome to creative effort. The principal monitors a noisy signal of the agent’s behavior

\[
(2) \quad y(e_1, e_2) = e_1 + e_2 + \epsilon_y,
\]

where \( \epsilon_y \) is a normal shock with mean zero and variance \( \text{Var}(\epsilon_y) = \sigma_y^2(m) \). The greater is the extent of monitoring \( m \), the more precise is the signal of behavior, \( d\sigma_y^2(m)/dm < 0 \), at a decreasing rate, \( d^2\sigma_y^2(m)/dm^2 \geq 0 \). The principal incurs the cost \( D(m) \) when the extent of monitoring is \( m \). The cost of monitoring is increasing and convex, \( dD(m)/dm > 0 \) and \( d^2D(m)/dm^2 \geq 0 \). Let \( \text{Cov}(\epsilon_x, \epsilon_y) = \sigma_{xy} > 0 \) denote the covariance between the outcome of
interest to the principal and the signal of the agent’s behavior, which is assumed to be positive. The objective of the principal is to maximize the expected outcome net of the agent’s compensation.

Outcome control corresponds to contracting on the outcome of interest to the principal, while behavior control corresponds to contracting on the signal of the agent’s behavior (Ouchi, 1979; Eisenhardt, 1985). We consider both forms of control, but in a fashion that is novel to the literature. First, we suppose the principal uses behavior control, thereby not providing monetary incentives. Thus, the principal solely compensates the agent on the basis of the signal of behavior, which the principal explicitly monitors. Second, we suppose the principal uses behavior control in conjunction with outcome control, thereby drafting a compensation contract contingent on the outcome (i.e., providing monetary incentives) as well as the signal of the agent’s behavior. Organizational theory formally considers either outcome or behavior control, not allowing for both as we do here. Nevertheless, as noted by Oliver and Anderson (1994, 1995), sales organizations usually incorporate aspects of both controls (i.e., hybrid governance).

The outcome of interest to the principal and the signal of the agent’s behavior (which is subject to monitoring) have the same sensitivity to routine effort. In this sense, the signal is a useful performance measure by which to evaluate the extent to which the agent is exerting routine effort. The implication is that if the agent were only exerting routine effort, then behavior control would be equivalent to outcome control. However, the outcome and signal of behavior have different sensitivities to creative effort. The premise is that creative effort is relatively more difficult to monitor; in other words, interpreted in the context of Banker and Datar (1989) and Holmstrom (1979), the signal of behavior is less informative about creative effort. The coefficient $\lambda$ represents the importance of creative effort in the generation of the outcome. Thus, the greater is $\lambda$, the less useful is behavior control (i.e., contracting on the signal of behavior, which the principal monitors), and thereby the more useful is outcome control (i.e., contracting on the outcome); that is, the less powerful is the signal of behavior as a mechanism by which to implement incentive compatibility with respect to creative effort.

The timing of the game is the following, which follows along the same lines as in Demougin and Fluet (2001). In the first stage, the principal sets the level of monitoring $m$ governing the signal of the agent’s behavior. In the second stage, the principal designs the compensation contract that satisfies incentive compatibility (to ensure the agent exerts the
desired levels of routine and creative effort) and individual rationality (to ensure the agent accepts the contract). With behavior control, the principal designs the contract to be contingent on the signal of the agent’s behavior; and with outcome control, the principal contracts on the outcome. We consider two scenarios: behavior control, and behavior control in conjunction with outcome control. In the third stage, the agent accepts or rejects the contract, and exerts routine and creative effort if he accepts the contract.

As in Dutta (2008), Datar et al. (2001), Holmstrom and Milgrom (1987), and Feltham and Xie (1994), the agent has exponential preferences with a (constant) coefficient of absolute risk aversion (CARA) of $R$. The reservation utility of the agent is $r$, which may be interpreted as the (certainty equivalent) income the agent would earn in the next best available employment opportunity.

**Behavior Control**

Suppose the principal uses behavior control, designing the compensation contract solely contingent on the signal of the agent’s behavior (which the principal monitors), not providing monetary incentives. As in Dutta (2008), Datar et al. (2001), Holmstrom and Milgrom (1987), and Feltham and Xie (1994), for tractability, we restrict our analysis to linear compensation contracts of the form

$$w_0(y(e_1,e_2)) = \alpha + \alpha_y y(e_1,e_2),$$

where $\alpha$ represents the agent’s salary and $\alpha_y$ the sensitivity of the agent’s compensation to the signal of behavior. The certainty equivalent (CE) of the agent is

$$CE[w_0(y(e_1,e_2))] = E[w_0(y(e_1,e_2))] - RVar[w_0(y(e_1,e_2))] / 2 - C(e_1, e_2).$$

Applying (2), this becomes

$$CE[w_0(y(e_1,e_2))] = \alpha + \alpha_y (e_1 + e_2) - R\alpha_y^2 \sigma^2(m) / 2 - c_1 e_1^2 / 2 - c_2 e_2^2 / 2.$$

Using backwards induction to solve the game, we first solve the principal’s contract design problem and then the principal’s monitoring problem (i.e., deciding the extent to which to monitor the agent’s signal of behavior). The principal’s contract design problem is

$$\max_{\alpha, \alpha_1, \alpha_2} E[x(e_1,e_2) - w_0(y(e_1,e_2))],$$

subject to the individual rationality (IR) constraint of the agent.
(6) \( CE[w_0(y(e_1, e_2))] \geq r \);

and the incentive compatibility (IC) constraint of the agent

(7) \( (e_1, e_2) = \arg \max_{(\hat{e}_1, \hat{e}_2)} CE[w_0(y(\hat{e}_1, \hat{e}_2))] \).

The following lemma derives the solution to the principal’s contract design problem.

**LEMMA 1:** Suppose the principal uses behavior control. Let \( m \) denote the level of monitoring that is chosen by the principal in the first stage of the game. The coefficient on the signal of the agent’s behavior is

\[
(8) \quad \alpha_x(m) = \frac{c_2 + \lambda c_i}{c_2 + c_1 + c_i c_2 R \sigma^2_x(m)}.
\]

The expected outcome \( X_0 \equiv E[x(e_1, e_2)] \) is

\[
(9) \quad X_0(m) = \frac{(c_2 + \lambda c_i)^2 / (c_1 c_2)}{c_2 + c_1 + c_i c_2 R \sigma^2_x(m)}.
\]

The expected profits of the principal \( \Pi_0 \equiv E[x(e_1, e_2) - w_0(y(e_1, e_2))] \) are

\[
(10) \quad \Pi_0(m) = \frac{(c_2 + \lambda c_i)^2 / (2c_1 c_2)}{c_2 + c_1 + c_i c_2 R \sigma^2_x(m)} - r.
\]

**Proof:** Please see the Appendix.

The following proposition shows that more extensive monitoring of the agent’s signal of behavior is beneficial to the principal in that it enhances the expected outcome and profits. The intuition is that greater monitoring increases the precision of the signal upon which compensation is based, thus it reduces the agency cost incurred by the manager associated with the moral hazard problem (Banker and Datar, 1989).

**PROPOSITION 1:** Suppose the principal uses behavior control. The expected outcome and profits of the principal are increasing in the extent to which the principal monitors the signal of the agent’s behavior, \( dX_0(m) / dm > 0 \) and \( d\Pi_0(m) / dm > 0 \), respectively.

**Proof:** Please see the Appendix.
We next solve the principal’s monitoring problem, which is given by

\[(11) \quad m_0 = \arg \max_{\{m\}} \Pi_0(m) - D(m) . \]

By virtue of the assumptions \( d\sigma_y^2(m) / dm < 0 \), \( d^2\sigma_y^2(m) / dm^2 \geq 0 \), \( dD(m) / dm > 0 \), and \( d^2D(m) / dm^2 \geq 0 \), the solution to this problem is unique and well-defined. The properties of the optimal level of monitoring are not of concern to us, except when compared to the scenario in which the principal uses behavior control in conjunction with outcome control.

**Behavior and Outcome Control**

Suppose the principal uses behavior and outcome control, such that the compensation contract is contingent on both the signal of the agent’s behavior (which is subject to monitoring by the principal) and the outcome of interest to the principal. As before, we restrict our analysis to linear compensation contracts of the form

\[(12) \quad w_t(y(e_1, e_2), x(e_1, e_2)) = \beta + \beta_y y(e_1, e_2) + \beta_x x(e_1, e_2), \]

where \( \beta \) represents the agent’s salary, \( \beta_y \) the sensitivity of the agent’s compensation to the signal of behavior (representing the extent of behavior control), and \( \beta_x \) the sensitivity of the agent’s compensation to the outcome (representing the extent of outcome control). The certainty equivalent (CE) of the agent is

\[(13) \quad CE[w_t(y(e_1, e_2), x(e_1, e_2))] = \beta + \beta_y (e_1 + e_2) + \beta_x (e_1 + \lambda e_2) \]

\[- R[\beta_y^2 \sigma_y^2(m) + \beta_x^2 \sigma_x^2 + 2\beta_y \beta_x \sigma_{xy}] / 2 - c_1 e_1^2 / 2 - c_2 e_2^2 / 2 . \]

We first solve the principal’s contract design problem and then the principal’s monitoring problem (i.e., deciding the extent to which to monitor the agent’s signal of behavior). The principal’s contract design problem is

\[(14) \quad \max_{\{\beta, \beta_y, \beta_x, \sigma_y, \sigma_x\}} E[x(e_1, e_2) - w_t(y(e_1, e_2), x(e_1, e_2))], \]

subject to the individual rationality (IR) constraint of the agent

\[(15) \quad CE[w_t(y(e_1, e_2), x(e_1, e_2))] \geq r ; \]

and the incentive compatibility constraint (IC) of the agent

\[(16) \quad (e_1, e_2) = \arg \max_{\{\hat{e}_1, \hat{e}_2\}} CE[w_t(y(\hat{e}_1, \hat{e}_2), x(\hat{e}_1, \hat{e}_2))]. \]

The following lemma derives the solution to the principal’s contract design problem.
Lemma 2: Suppose the principal uses behavior control in conjunction with outcome control. Let \( m \) denote the level of monitoring that is chosen by the principal in the first stage of the game. The coefficient on the signal of the agent’s behavior is

\[
\beta_y(m) = \frac{(c_2 + \lambda^2 c_1)(c_1 c_2 R \sigma^2_{xy} + c_2 + \lambda c_1) - (c_2 + \lambda c_1)(c_2 + \lambda^2 c_1 + c_1 c_2 R \sigma^2_{y})}{(c_1 c_2 R \sigma^2_{xy} + c_2 + \lambda c_1)^2} - (c_1 c_2 R \sigma^2_{xy} + c_2 + \lambda^2 c_1)(c_2 + c_1 + c_1 c_2 R \sigma^2_{y}(m)) ;
\]

and the coefficient on the outcome is

\[
\beta_x(m) = \frac{(c_2 + \lambda c_1)(c_1 c_2 R \sigma^2_{xy} + c_2 + \lambda c_1) - (c_2 + \lambda^2 c_1)(c_2 + c_1 + c_1 c_2 R \sigma^2_{y}(m))}{(c_1 c_2 R \sigma^2_{xy} + c_2 + \lambda c_1)^2} - (c_1 c_2 R \sigma^2_{xy} + c_2 + \lambda^2 c_1)(c_2 + c_1 + c_1 c_2 R \sigma^2_{y}(m)) .
\]

The expected outcome \( X_1 = E[x(e_1, e_2)] \) is

\[
(19) \quad X_1(m) = [\beta_y(m)(c_2 + \lambda c_1) + \beta_x(m)(c_2 + \lambda^2 c_1)]/(c_1 c_2).
\]

The expected profits of the principal \( \Pi_1 = E[x(e_1, e_2) - w_y(y(e_1, e_2), x(e_1, e_2))] \) are

\[
(20) \quad \Pi_1(m) = [\beta_y(m)(c_2 + \lambda c_1) + \beta_x(m)(c_2 + \lambda^2 c_1)]/(2c_1 c_2) - r.
\]

Proof: Please see the Appendix.

The following proposition shows that, as in the case in which the principal solely uses behavior control, more extensive monitoring of the agent’s signal of behavior is beneficial to the principal. The intuition is the same: greater monitoring increases the precision of the signal upon which compensation is partially based, thus it reduces the agency cost incurred by the principal associated with the moral hazard problem (Banker and Datar, 1989).

Proposition 2: Suppose the principal uses behavior control in conjunction with outcome control. The expected outcome and profits of the principal are increasing in the extent to which the principal monitors the signal of the agent’s behavior, \( dX_1(m)/dm > 0 \) and \( d\Pi_1(m)/dm > 0 \).

Proof: Please see the Appendix.

We next solve the principal’s monitoring problem, which is given by

\[
(21) \quad m = \arg \max_{[m]} \Pi_1(m) - D(m) ,
\]
the solution to which is unique and well-defined. The next sub-section compares the solution to this problem to that which arises when the principal solely uses behavior control.

**Comparing Compensation Contracts With versus Without Outcome Control**

We first show that, as one would expect, the principal is better off using behavior control in conjunction with outcome control (rather than behavior control by itself). Indeed, invoking the informativeness principal of Holmstrom (1979), the principal cannot be made worse off by having another informative signal (in this case, the outcome) upon which to base compensation.

**PROPOSITION 3:** Hold constant at $m$ the level of monitoring that is chosen by the principal in the first stage of the game. The expected outcome and profits of the principal are greater when behavior control is used in conjunction with outcome control, $X_1(m) \geq X_0(m)$ and $\Pi_1(m) \geq \Pi_0(m)$.

**Proof:** Please see the Appendix.

The following proposition shows that, if the outcome is sufficiently sensitive to creative effort, then the marginal benefit to the principal (in terms of the expected outcome and profits) of increased monitoring diminishes when the principal uses outcome and behavior control combined.

**PROPOSITION 4:** Hold constant at $m$ the level of monitoring that is chosen by the principal in the first stage of the game. Suppose the sensitivity of the outcome to creative effort is sufficiently great to satisfy

\[ \lambda \geq \frac{2c_1c_2 + c_1^2c_2RCov + \sqrt{[2c_1c_2 + c_1^2c_2RCov]^2 - 4c_1[c_2 + c_1c_2R\sigma^2(m)][c_2 + c_1c_2R\sigma^2(m)] - c_1^2c_2RCov}}{2c_1[c_2 + c_1c_2R\sigma^2(m)]} \]

The marginal benefit to the principal of monitoring is lower when the principal uses behavior control in conjunction with outcome control; similarly, the value to the principal of using outcome control when paired with behavior control diminishes as the principal increases the level of monitoring, $d(X_1(m) - X_0(m))/dm \leq 0$ and $d(\Pi_1(m) - \Pi_0(m))/dm \leq 0$. 

12
Proof: Please see the Appendix.

Propositions 1 and 2 demonstrated that enhanced monitoring of the agent’s signal of behavior increases the expected outcome and profits of the principal with or without outcome control. However, the marginal benefit to the principal of enhanced monitoring differs under each form of control because the outcome and signal of the agent’s behavior have different sensitivities to creative effort, namely $\lambda$ and 1, respectively. The larger is the coefficient $\lambda$, the greater is the extent to which the signal undervalues the agent’s exertion of creative effort; thus, the greater is the degree of incongruence in using the signal of the agent’s behavior as a device by which to elicit a favorable outcome by implementing high effort. If the principal uses outcome control (i.e., implementing a contract that is contingent on the outcome), then the larger is $\lambda$, the smaller is the relative value of the signal as an incentive device and by extension the marginal benefit of enhanced monitoring. By contrast, this is not the case in the absence of outcome control since the signal of the agent’s behavior is the only means by which to elicit incentive compatibility (so as to resolve the moral hazard problem). We infer that if $\lambda$ is sufficiently large, then the marginal benefit to the principal of increased monitoring is smaller when outcome control is used (in conjunction with behavior control).

Because the order of differentiation (in calculus) does not matter, our finding may also be interpreted as stating that the value to the principal of using outcome control (in combination with behavior control) diminishes as the level of monitoring increases.

The next proposition follows directly from the previous one, demonstrating that because the marginal benefit to the principal of increased monitoring is lower when the principal uses outcome control (in combination with behavior control), the optimal level of monitoring is greater when the principal solely uses behavior control (given that the signal of the agent’s behavior is subject to monitoring).

**Proposition 5:** Suppose the sensitivity of the outcome to creative effort is sufficiently great to satisfy (22). The optimal level of monitoring is greater when the principal does not use outcome control, $m_0 \geq m_f$.

Proof: Please see the Appendix.
III. LITERATURE REVIEW AND EMPIRICAL HYPOTHESES

Following in the tradition of Anderson and Oliver (1987), Oliver and Anderson (1994, 1995), and Eisenhardt (1985, 1988), the effect of supervisory monitoring and the impact of the implementation of a monetary incentives scheme on retail sales productivity provide an appropriate empirical context within which to test the predictions of our analytical model. We interpret the outcome of interest to the principal as retail sales productivity, such that the goal of the principal (i.e., the store manager or supervisor) is to maximize the retail sales productivity of the salesperson (i.e., the agent) net of the agent’s compensation. Outcome control corresponds to offering the salesperson monetary incentives (e.g., a sales commission). Behavior control corresponds to observing how hard the salesperson is working towards achieving tasks such as customer satisfaction; in the context of our analytical model, we labeled this component the (noisy) signal of the agent’s behavior, which is subject to monitoring by the supervisor. The salesperson exerts two types of effort: routine effort is more readily measurable and quantifiable; while creative effort is more difficult to monitor, yet it may have a (possibly more) significant effect on retail sales productivity. In the context of the framework of our analytical model, due to the fact that the signal of the agent’s behavior does not accurately capture the sensitivity of the outcome to creative effort, routine effort is more responsive to behavior control, while outcome control is more effective in eliciting creative effort.

According to organizational control theory, the choice between behavior and outcome control should be a function of the specific task characteristics, namely, outcome measurability and task programmability (Eisenhardt, 1985). That is, if the task is less programmed and outcomes are readily measured, outcome control is the appropriate strategy. These conditions seem to fit the task characteristics of salespeople. As Siguaw, Brown, and Wilding note, “customer orientation is a selling behavior in which salespersons assist customers in making purchase decisions that will satisfy long-term wants and needs” (1994, p. 108). It is heterogeneity of customers’ needs and wants that makes salespeople’s tasks difficult to prescribe and thereby less programmable (Anderson and Oliver, 1987; Eisenhardt, 1985, 1988). It is often necessary for salespeople to alter their interactions and sales behaviors across customers (Weitz et al., 1986). Thus, although the outcomes are measurable, the degree of programmability is fairly small. Furthermore, when a monetary incentives scheme is implemented to support a customer-focused service strategy, the effect of monetary incentives on organizational
performance has not been uniform across organizations providing customer-focused service (Gerhardt et al., 1994).

Anderson and Oliver (1987, p. 76) conceptualize outcome and behavior control systems as two extremes. In their later works, Oliver and Anderson (1994, 1995) offer further empirical findings and more detailed analysis of sales control systems, and point out that sales organizations usually position the salesforce control system somewhere between these two poles, incorporating aspects of both controls (i.e., hybrid governance). Oliver and Anderson (1995) find that firms employing a hybrid control system place a very high level of emphasis on supervision and quantitative results in evaluating salespeople’s performance. They show that these firms have a neutral philosophy on the relative importance of managing salespeople’s outcomes over behaviors. More importantly, they argue that sales organizations should be able to respond to the motivational needs and experience of individual salespeople, resulting in a form of control system segmentation. Hence, sales organizations must find their own equilibrium. Oliver and Anderson conclude, “Further study is required to determine if other variations in control system philosophy exist and what the consequences of such systems may be. Study should also proceed on variations of the hybrid system which optimizes the sales impact of rep organizations generally” (1995, p. 15). Our analytical model provides a natural context within which to respond to their suggestion. In that spirit, to test our model, we empirically examine the effect of the interaction between the level of supervisory monitoring and monetary incentives on organizational performance in a retail setting.

Agency theory suggests monetary incentives help resolve moral hazard and adverse selection problems. A principal designs a contract to motivate a risk- and effort-averse agent to exert unobservable effort in a production process that is characterized by uncertainty (Jensen and Meckling, 1976; Holmstrom, 1979; Banker and Datar, 1989; Bushman and Indjejikian, 1993; Feltham and Xie, 1994; Datar et al., 2001). If more productive employees have superior outside opportunities, implementation of a monetary incentives scheme amounts to offering a menu of contracts to enable employees to determine their compensation by how hard they choose to work (Stiglitz, 1977; Darrough and Melumad, 1995). In part, a contract is designed to reveal the truth about the agent’s unobservable intrinsic ability (Murphy, 1986; Harris and Raviv, 1978; Rothschild and Stiglitz, 1976; Spence, 1973; Salop and Salop, 1976; Wilson, 1977). Furthermore, monetary incentives tend to attract and retain the most productive job applicants and to
discourage the least productive, to the employer’s benefit resulting in a sorting and self-selection effect (Milgrom and Roberts, 1992, p. 157). Thus, monetary incentives attract workers of higher ability and induce workers to provide higher effort (Booth and Frank, 1999).

Monetary incentives schemes may take several forms such as merit pay, commissions, and bonuses. Previous studies in management accounting, organizational behavior, and human resource management have extensively examined the impact of individual-level incentive schemes on organizational performance; however, their findings have produced conflicting results (Challagalla and Shervani, 1996; Gerhardt and Milkovich, 1990, 1992). Researchers have reported that the performance impact of individual-level monetary incentives schemes is positive (Banker et al., 1996a, 1996b), negative (Challagalla and Shervani, 1996), modest (Kerr, 1999; Lawler, 1990; Pfeffer, 1995; Stajkovic and Luthans, 2001), and insignificant (Cravens et al., 1993). Hence, because the true value of individual-level monetary incentives remains unclear, there remains a need for further evidence on the subject.

Our analytical model provides formal guidance in this regard. We showed in Proposition 3 that a principal cannot be made worse off by having another tool upon which to base the agent’s compensation, such that the provision of monetary incentives (outcome control) should improve performance and at worst have no effect. We thereby postulate the following hypothesis:

H1: The effect of monetary incentives on retail sales productivity is positive.

Agency theory argues that if an agent’s actions can be monitored more precisely, desired actions can be induced with lower risk premium costs, which in turn leads to an improvement in organizational productivity (Lambert, 2001). In the marketing literature, agency theoretic research has examined the role of monitoring of a salesperson’s effort to alleviate moral hazard problems (Basu et al., 1985; Lal and Srinivasan, 1993; Joseph and Thevaranjan, 1998; Slater and Olson, 2000). Monitoring provides an imperfect signal of the salesperson’s effort, and compensating the salesperson based on this signal induces higher effort. This imperfect monitoring signal can be interpreted similar to behavior control which posits that supervisors who “have a well-defined idea of what they want salespeople to do” can work to ensure that the salesforce behaves accordingly (Anderson and Oliver, 1987, p. 77). Therefore, retailers are likely
to exploit the monitoring role that supervisors play in enhancing the productivity of their retail selling activities when they do not implement outcome control.

Our analytical model formalized these arguments. We adopted the view that behavior control is akin to regulating an imperfect monitoring signal, and outcome control is equivalent to implementing a monetary incentives scheme. We proved in Propositions 1 and 2 that more extensive monitoring is beneficial to the principal as it increases the precision of the signal upon which compensation is partially based, thereby reducing the agency cost incurred by the principal associated with the moral hazard problem, in accord with the precision and sensitivity results in Banker and Datar (1989) and in agreement with the precepts in Lambert (2001). We thus specify the following hypothesis:

H2: The effect of monitoring on retail sales productivity is positive.

However, there has been little empirical evidence to assess whether this agency theoretic prediction holds for retail outlets in which a monetary incentives scheme is introduced so that salespeople are required to provide customer-focused service that involves greater worker empowerment and task ambiguity. As illustrated by our analytical model, this is especially important since monitoring of ambiguous tasks may not provide more informative signals. Moreover, in this retail setting, a higher level of supervisory monitoring may cause confusion and communication problems on the part of salespeople who are supposed to be creative in providing customer-focused service.

Organizational control theory suggests that behavior control requiring a high level of supervisory monitoring is not appropriate in an environment characterized by low task programmability (Anderson and Oliver, 1987; Eisenhardt, 1985, 1988; Ouchi, 1979). In retail organizations that promote customer-focused service, workers involved in customer service need to be empowered because the exact tasks required to improve customer satisfaction cannot be pre-specified, as different customers have different needs, and their service expectations often differ from those of management (Schlesinger and Heskett, 1991). Specifically, customer focus is more effective when customers are particularly discerning or demanding of high levels of service. In this context, behavior control requiring a high level of supervisory monitoring is likely to constrain salespeople from exploring creative new ways to provide higher levels of
customer service. Therefore, in retail organizations with an emphasis on customer service, increasing the level of supervisory monitoring in the presence of monetary incentives is less likely to improve retail sales productivity.

A large body of literature on employee creativity has examined the possibility that organizational characteristics contribute significantly to the creative performance of employees. Oldham and Cummings (1996) find that participants produce the most creative work when they have appropriate creativity-relevant characteristics, work on complex, challenging jobs, and are supervised in a supportive, non-controlling fashion. Specifically, they identify the conditions under which the style of supervision promotes creativity among employees. In particular, the authors argue that supervision that is supportive of employees should enhance creative achievement. However, when supervisors are controlling in monitoring employee behavior, make decisions without employee involvement, provide feedback in a controlling manner, and generally pressure employees, they undermine intrinsic motivation and shift the employees’ focus of attention away from work activities and toward external concerns (Deci et al., 1989; Deci and Ryan, 1987). This reduction in intrinsic motivation lowers creative performance in an organization characterized by complex and challenging tasks, and since our research site emphasizes customer-focused service, supervisory monitoring is likely to make salespeople less productive. In a similar vein, Zhou (2003) demonstrates that when creative coworkers are present, the weaker is the degree to which supervisors engage in close monitoring, the more the employees exhibit creativity.

Our agency theory models these tradeoffs. We showed in Proposition 4 that the marginal benefit to the principal of enhanced monitoring (of the signal of the agent’s behavior) diminishes when monetary incentives are provided if creative effort is sufficiently important in the determination of the outcome. We obtain this finding since an increase in the sensitivity of retail sales productivity to creative effort reduces the usefulness of the signal of the agent’s behavior and by extension monitoring as a device by which to implement incentive compatibility (with respect to routine and creative effort). In light of our discussion above about the relatively large contribution of creativity in customer-focused organizations, such as our research site, we postulate that this condition holds. Furthermore, our finding may also be interpreted as stating that the marginal benefit to the principal of providing monetary incentives diminishes in the presence of monitoring. An extension of this result is that the optimal level of monitoring is
smaller when the sales manager provides monetary incentives in conjunction with utilizing behavior control (Proposition 5). Therefore, we specify the following hypotheses:

H3: The marginal effect of monitoring on retail sales productivity is reduced when a monetary incentives scheme is provided.

H4: The value of a monetary incentives scheme decreases as the level of supervisory monitoring increases.

H5: The level of monitoring is lower when a monetary incentives scheme is provided.

IV. DATA AND ESTIMATION METHOD

Research Site

Retailers can be broadly classified into three distinct market segments: high-end retailers like Neiman Marcus, retailers positioned in the middle such as Foley’s, and low-end retailers like Kmart (Anderson, 1999). Our research site falls in the category of a fair-priced retailer positioned in the middle of the department store industry. The mass retailing industry has seen significant changes in the last three decades. Specialty retailers such as The Gap, Sunglass Hut, and Sharper Image that focus on one type of general merchandise, as well as category killers like Best Buy, OfficeMax, and Lowes, that dominate a particular category of goods by buying huge volumes and selling at a discount, have steadily increased their market shares. Moreover, the continuing rise of discount stores and the recent ascent of Internet retailers mean that traditional retailers have been hard pushed to maintain the level of growth required to sustain their operations. Many fair-priced department stores positioned in the middle seem stuck in a no-win situation. Price-conscious shoppers bypass them for discounters, while brand-hungry consumers head to the high-end and other specialty stores for the top brand names. As square footage in the retail industry has increased, so have the number of Chapter 11 bankruptcies (such as Montgomery Ward Holding) and the number of store closings (notably U.S. Woolworth stores). This has led to increasing pressure on existing retailers such as our research site to develop new strategies to survive in this competitive environment.

The strategic positioning at our research site is to command medium-range prices for its merchandise by providing a reasonable level of service to its customers, going beyond simply ringing up the cash register or responding to customer requests. In order to achieve its strategic
position, our research site introduced a monetary incentives scheme to provide service beyond customer expectations. This scheme was expected to lead to superior customer service and thereby improved customer satisfaction and sales performance.

Customer-focused service under the monetary incentives scheme at our research site involves understanding and satisfying individual needs, which differ widely across customers. Therefore, the tasks performed by its salespeople are more challenging, more difficult to prescribe, and less programmable than the tasks required for a more conventional, mass-production-style service at a low-end department or discount store. The main role of supervisors at a low-end retail outlet is to formulate a well-defined plan of what the salespeople should do and to monitor them closely to ensure that they comply with the prescribed activities. However, at a retail outlet positioned in the middle, the role of supervisors is more ambiguous in supporting salespeople.

Our research site, a chain of department stores, implemented a monetary incentives scheme that pays a commission on sales. It was part of an effort to inject new life into a traditionally managed chain that had enjoyed unusually high market share for much of its history and whose sales and profits have stagnated. This monetary incentives scheme involves careful selection of sales personnel, with heavy reliance on referrals from associates and on training and recognition. Under this scheme, employees were trained in customer service, with an emphasis on providing a high quality of customer service. In addition, employees were endowed with greater authority to close a sale.

The monetary incentives scheme was implemented in stores sequentially as company managers were unsure about the precise extent of the impact of the plan on the productivity of those stores. This sequential nature of the implementation process provides us with 10 experimental stores whose impact can be compared against 10 control stores that did not implement the monetary incentives scheme during our sample period.

Data and Variables

We obtained monthly data spanning a 60-month period for 10 experimental stores of a retail chain that implemented the monetary incentives scheme and 10 control stores of the same retail chain in the same geographic region that did not. Each individual store represents a decision-making unit (DMU). In our specification, we model the production function relating the
output of each DMU as a function of its inputs, such as labor and capital, and contextual variables including the level of supervisory monitoring and the implementation of the monetary incentives scheme. There has been a substantial debate in the past about the appropriateness of an output measure for a retailer (Achabad et al., 1984; Goodman, 1985; Thurik and Kooiman, 1986). Achabal et al. (1984) argue that output measures such as sales are inadequate as they fail to capture the influence of the local level of demand and competition. Their key argument is that output measures should be based upon measures of ability to produce and not the amount produced. In essence, their argument is that contextual variables influencing the ability of a DMU to produce more sales need to be included as control variables. In contrast, Goodman (1985) suggests that sales should be the only output measure used because the ability to produce without considering actual sales does not identify ineffective use of inputs. Supporting this argument, Thurik and Kooiman (1986) find that it is the supply and not the demand that constrains output. In our empirical setting, we measure output as monthly sales (SALES) in deflated dollars. However, sales alone may not capture the strategies pursued by retail outlets, which necessitates the use of multiple variables to measure constructs (Venkatraman and Ramanujam, 1986; Chakravarthy, 1986; Lewin and Minton, 1986).

Recent research has included store size, labor usage, and capital investment as inputs in the production function (Nooteboom, 1983; Ingene, 1982, 1985; Hise et al., 1983; Good, 1984; Lusch and Moon, 1984; Doutt, 1984; Ratchford and Brown, 1985; Thurik and Kooiman, 1986; Kamakura et al., 1996). This appears to answer the call by Samiee (1990) to focus less on labor productivity and more on what allows managers to make real advances in productivity. Therefore, to capture the labor input in the production function, we utilize the number of selling hours in each store per month (HOURS). The two major forms of capital investments in the retail setting are the selling space utilized and the merchandise carried by the store. We include two capital variables in the production function: the size of the store in square feet (SIZE) and the dollar value of the average inventory of merchandise (INVENTORY). INVENTORY is calculated as the mean of the opening and closing inventories carried by the store for a given month.

We also include contextual variables that may be exogenously fixed as well as others that may be under the control of the DMU managers. In the retail business, the manager of the DMU has no control over the area in which each store is located, the demographics of the store location in terms of median household income, median age, percentage of blue collar workers, median
family size, percentage of the population with college education, size of county population, and the intensity of the competition faced by each retail store. To capture the differences in the location of the stores, we include an indicator variable RURAL whose value is 1 if the store is located in a rural area, and 0 otherwise. To account for differences in the demographics of the store locations, we include INCOME to represent the median household income, POPLNAGE to reflect the median age of the county’s population, BLUECOLLAR to account for the percentage of blue collar workers, FAMILYSIZE to measure the median family size, COLLEGE to reflect the percentage of the population with a college education, and POPULATION to control for the total population size in a specific geographical area. Stores in upscale and less heavily populated rural markets are likely to enjoy higher productivity since upscale customers are attracted by enhanced customer service more than other customers (Peterson et al., 1989; Banker et al., 1996b). Specifically, retail sales productivity is likely to be higher for stores located in those regions where customers have higher household incomes, there is a higher proportion of older households with greater wealth, the family size is smaller, and the proportion of better-educated customers is higher. To capture the differences in the intensity of the competitive environment, we include the index COMPETITION constructed at our research site to measure the number and quality of competitors. Stores in a more competitive environment are likely to be less productive in generating retail sales.

The contextual variables included to evaluate our hypotheses are the level of supervisory monitoring (MONITORING), constructed as the ratio of the number of managerial hours to the number of selling hours at each month-store; MONINCENT that is 1 if a store was under the monetary incentives scheme in a month, and 0 otherwise; and MONINCENT*MONITORING to capture the interaction effect of the MONITORING variable for stores that implemented the monetary incentives scheme. Furthermore, we include average monthly sales for all 20 stores (AVGSALES) to control for economy-wide and industry-wide effects; and an indicator variable (SEASON) whose value is 1 during the holiday sales season spanning October, November, and December, and 0 otherwise, to control for the seasonal nature of the retail business.\footnote{The sales in the three months of the holiday shopping season comprise about 40% of total annual sales at our research site.} The relationship between inputs, output, and contextual variables is depicted in Figure 1.

[Insert Figure 1]
All variables measured in dollars are deflated by the Department Store Inventory Price Index calculated by the Bureau of Labor Statistics. The variables are deflated in order to make them comparable over different months.

**Descriptive Statistics**

Table 1 provides descriptive statistics of output, inputs, and the contextual variables. The median values of the output variable (SALES) and the three input variables (HOURS, SIZE, and INVENTORY) are all smaller than their mean values, indicating that the data are skewed to the right. Other than MONINCENT, the contextual variable representing our principal hypotheses is MONITORING, which has a mean of 8.73% and median of 8.52%.

[Insert Table 1]

We matched 10 experimental stores that implemented the monetary incentives scheme with 10 control stores that did not. Table 2 compares the 10 experimental stores and the 10 control stores in terms of their levels of output, inputs, and contextual variables before the introduction of monetary incentives scheme. Paired sample t, sign, and sign rank tests show that there are no statistically significant differences in output, inputs, and contextual variables between the 10 experimental stores and the 10 control stores.

[Insert Table 2]

Table 3 compares the 10 experimental stores and the 10 control stores in terms of their levels of output, inputs, and contextual variables after the introduction of the monetary incentives scheme. Paired sample t, sign, and sign rank tests indicate that the level of supervisory monitoring (MONITORING) is lower for the 10 experimental stores; and total selling hours (HOURS) is greater for the 10 experimental stores. Therefore, in accord with Hypothesis 5, the optimal level of monitoring is lower when monetary incentives are provided.

[Insert Table 3]

**Estimation Models**

We implement the two-stage estimation procedure in Banker and Natarajan (2008) as follows. In the first stage, we use the Data Envelopment Analysis (DEA) model of Banker et al. (1984), hereafter referred to as BCC, to evaluate the productivity of each DMU-month observation \((j, t)\), for \(j = 1,\ldots,20\) and \(t = 1,\ldots,60\). We use the output-oriented BCC model to
estimate the productivity scores using a set of three inputs to produce one input. Our single output measure is monthly sales in deflated dollars (SALES), and our three input measures are the number of selling hours (HOURS), the size of the store in square feet (SIZE), and the dollar value of the average merchandise inventory (INVENTORY). There are 1,200 (=20×60) observations. The formulation of this output-oriented BCC model for estimating the productivity \( \hat{\theta}_{jt} \) of an observation is given by the following linear program that is solved for each observation \((j, t)\):

\[
\begin{align*}
\max_{\lambda_j} & \quad \lambda_j \chi_{ij} \\
\text{s.t.} & \quad \sum_{j=1}^{20} \sum_{t=1}^{60} \lambda_j y_{jt} + \theta_j y_{jt} & \leq 0 \\
& \quad \sum_{j=1}^{20} \sum_{t=1}^{60} \lambda_j = \theta_j \geq 0
\end{align*}
\]

where

\( \chi_{ijt} \) = quantity of input \( i \) consumed by DMU\(_{jt} \);  
\( y_{jt} \) = quantity of output produced by DMU\(_{jt} \);  
\( \lambda_{jt} \) = weight placed on inputs/outputs of DMU\(_{jt} \);  
\( i=1,\ldots,3; \ j=1,\ldots,20; \ t=1,\ldots,60. \)

Table 4 reports the DEA productivity scores of our pooled sample of 1,200 observations computed using the BCC model in (23). The productivity scores range from 0.2036 to 1.0000 with a higher score indicating greater productivity. The interquartile range for the productivity scores is from 0.4155 to 0.5893. The mean of the DEA productivity scores is 0.5245 and the median is 0.4936. Out of the 1,200 observations, 26 observations are on the production frontier.

[Insert Table 4]

In the second stage, to evaluate our hypotheses about the impact of supervisory monitoring and monetary incentives on the productivity of each store, we regress the logarithm of productivity \( \hat{\theta}_{jt} \) on the contextual variables using the full panel of pooled data. Banker and
Natarajan (2008) show that this two-stage procedure involving nonparametric estimation of efficiency in the first stage followed by OLS regression provides statistically consistent estimates in the second stage. Specifically, we estimate the regression

\[
\ln \hat{\theta}_{jt} = \beta_0 + \beta_1 \text{MONITORING}_{jt} + \beta_2 \text{MONINCENT}_{jt} \\
+ \beta_3 \text{MONINCENT}_{jt} \cdot \text{MONITORING}_{jt} + \beta_4 \text{AVGSALES}_t \\
+ \beta_5 \text{SEASON}_t + \beta_6 \text{INCOME}_j + \beta_7 \text{POPLNAGE}_j + \beta_8 \text{BLUECOLLAR}_j \\
+ \beta_9 \text{FAMILYSIZE}_j + \beta_{10} \text{COLLEGE}_j + \beta_{11} \text{POPULATION}_j + \beta_{12} \text{RURAL}_j \\
+ \beta_{13} \text{COMPETITION}_j + \epsilon_{jt}
\]

where

\[\ln \hat{\theta}_{jt}\] is the logarithm of the productivity of store \(j\) in month \(t\);

\(\text{MONITORING}_{jt}\) = the ratio of the number of managerial hours to the number of total selling hours for store \(j\) in month \(t\);

\(\text{MONINCENT}_{jt} = 1\) if store \(j\) was under the monetary incentives scheme in month \(t\), and \(0\) otherwise;

\(\text{AVGSALES}_t\) = monthly deflated average sales for all 20 stores;

\(\text{SEASON}_t = 1\) if the month \(t\) is October, November, or December, and \(0\) otherwise;

\(\text{INCOME}_j\) = median household income for the county in which store \(j\) is located;

\(\text{POPLNAGE}_j\) = median age of the population for the county in which store \(j\) is located;

\(\text{BLUECOLLAR}_j\) = the percentage of blue collar workers for the county in which store \(j\) is located;

\(\text{FAMILYSIZE}_j = \) median family size for the county in which store \(j\) is located;

\(\text{COLLEGE}_j = \) the percentage of the population with a college education for the county in which store \(j\) is located;

\(\text{POPULATION}_j = \) population size for the county in which store \(j\) is located;

\(\text{RURAL}_j = 1\) if store \(j\) is located in a rural area, and \(0\) otherwise;

\(\text{COMPETITION}_j = \) the competition intensity of store \(j\)’s market;

\(j = 1, \ldots, 20; t = 1, \ldots, 60\).

As a robustness check, we also estimate the following change model to test our hypotheses:
\[\Delta \ln \hat{\theta}_{jt} = \alpha_0 + \alpha_1 \Delta \text{MONITORING}_{jt} + \alpha_2 \text{MONINCENT}_{jt}\]

\[+ \alpha_3 \text{MONINCENT}_{jt} \Delta \text{MONITORING}_{jt} + \alpha_4 \Delta \text{AVGEFFICIENCY}_t\]

\[+ \alpha_5 \Delta \text{SALESGROWTH}_{jt} + \alpha_6 \Delta \text{AVGSALES}_t + \varepsilon_{jt}\]

where

\[\Delta \ln \hat{\theta}_{jt} = \text{change in the logarithm of the productivity of store j from month } t \text{ to } t-1;\]

\[\Delta \text{MONITORING}_{jt} = \text{change in the ratio of the number of managerial hours to the number of total selling hours for store j from month } t \text{ to } t-1;\]

\[\text{MONINCENT}_{jt} = 1 \text{ if store j was under the monetary incentives scheme in month } t, \text{ and 0 otherwise;}\]

\[\Delta \text{AVGEFFICIENCY}_t = \text{change in the monthly average of the logarithm of productivity for all 20 stores from month } t \text{ to } t-1;\]

\[\Delta \text{SALESGROWTH}_{jt} = \text{change in sales growth for store j from month } t \text{ to } t-1,\]

\[\Delta \text{AVGSALES}_t = \text{change in average monthly sales for 20 stores from month } t \text{ to } t-1.\]

Table 5 reports the Pearson and Spearman correlation coefficients between the dependent and independent variables from the second stage analysis. Before controlling for the impact of the contextual variables, there is a statistically insignificant correlation between \(\Delta \ln \hat{\theta}_{jt}\) and MONITORING; a statistically significant negative correlation between \(\Delta \ln \hat{\theta}_{jt}\) and MONINCENT; and a statistically significant negative correlation between \(\Delta \ln \hat{\theta}_{jt}\) and MONINCENT*MONITORING.

[Insert Table 5]

V. EMPIRICAL RESULTS

Following Banker and Natarajan (2008), in the second stage of our empirical analysis, we regress the logarithm of DEA productivity scores on the contextual variables. Because pooled cross-sectional and time-series information is used to estimate the impact of contextual variables on retail sales productivity, there is the potential for serial correlation biasing the standard errors of the coefficients. Therefore, we performed specification tests for residuals to check for serial correlation in the estimation model (24) and found that there exists substantial positive serial
correlation (parameter estimate = 0.451, t = 17.11). We address this problem by using a variant of the Prais-Winsten (1954) estimator proposed by Park and Mitchell (1980) to make first-order autocorrelation adjustments to the variables. This estimator is consistent and performs especially well for short time series and trended data in relation to several other estimates (Doran and Griffiths, 1983). It also reduces the extent to which the serial correlation coefficient tends to be underestimated by simpler methods (Kmenta and Gilbert, 1970). We test our hypotheses using the parameter estimates from the regression using the transformed variables. In the same vein, in the estimation model (25) that is in terms of the change variables, we found that there exists substantial negative serial correlation (parameter estimate = -0.436, t = -5.95). We also address this problem by using a variant of the Prais-Winsten (1954) estimator proposed by Park and Mitchell (1980) to make first-order autocorrelation adjustments to the variables.

We present the empirical results for the estimation model (24) in Table 6. The coefficient $\beta_1$ captures the main effect of the contextual variable MONITORING. The coefficient $\beta_2$ represents the main effect of implementing a monetary incentives scheme on retail sales productivity after controlling for the contextual factors. The interaction effect of the contextual variable MONITORING for stores that introduced monetary incentives is captured by the variable crossed with the dummy variable for monetary incentives scheme implementation. That is, the coefficient $\beta_3$ represents the difference in the effect of monitoring on retail sales productivity between the 10 experimental stores that implemented the monetary incentives scheme and the 10 control stores that did not.

The estimated coefficient on MONINCENT is positive and significant at the 1% level, indicating that implementing a monetary incentives scheme enhances retail sales productivity, supporting Hypothesis 1. The estimated coefficient on MONITORING is positive and significant at the 1% level, indicating that monitoring is positively associated with retail sales productivity for all 20 stores, supporting Hypothesis 2. The estimated coefficient on MONINCENT*MONITORING is negative and significant at the 1% level, suggesting that the marginal benefit of monitoring is diminished when monetary incentives are in place, supporting Hypothesis 3. The sum of $\beta_2$ and $\beta_3$*MONITORING (i.e., $\beta_2 + \beta_3$*MONITORING) decreases as MONITORING increases. Specifically, when MONITORING is greater than or equal to the first decile of MONITORING, $\beta_2 + \beta_3$*MONITORING has a negative value. When MONITORING is greater than or equal to the mean or median of MONITORING, $\beta_2 + \beta_3$*MONITORING has a
statistically significant negative value. This suggests that the value of monetary incentives decreases as the level of supervisory monitoring increases, which supports Hypothesis 4. Therefore, after controlling for economy-wide effects, the differences in the location of each store, the demographics of the store locations, and the intensity of the competitive environment, the results provide strong support for our main hypotheses that supervisory monitoring and monetary incentives have a positive impact on retail sales productivity and the marginal impact of monitoring on sales productivity is lower when a monetary incentives scheme is implemented. These results are confirmed by the change model specified in (25), though monetary incentives are no longer significant.

[Insert Table 6]

[Insert Table 7]

In Table 6, the coefficients on all the control variables, except for COLLEGE and POPULATION, are statistically significant and generally agree with their expected signs. The following have significant positive effects on retail sales productivity: average sales in all 20 stores (AVGSALES) to account for the macroeconomic business cycle, the holiday season (SEASON), the income of nearby residents (INCOME), the proportion of blue collar workers (BLUECOLLAR), and whether the store is in a rural area (RURAL). The following have significant negative effects on retail sales productivity: the age of the local population (POPLNAGE), the size of nearby families (FAMILYSIZE), and the extent of competition (COMPETITION).

VI. CONCLUSION

This paper proposed a principal-agent model with moral hazard to examine the tension between outcome control (in the form of the provision of monetary incentives), behavior control, and the role of monitoring, along with the contribution of each. Our model has two forms of effort, routine and creative; the outcome and signal of the agent’s behavior depend on both forms of effort; and the principal chooses the extent of monitoring that increases the precision of the signal. Under the premise that creative effort is more difficult to monitor than routine effort, we showed that the marginal impact of monitoring on the expected outcome and profits of the principal is reduced when outcome control is used if creative effort is sufficiently important in the generation of the outcome.
To test the model, we investigated how the level of supervisory monitoring and monetary incentives affect retail sales productivity using a panel of 60 month data for 10 experimental stores and 10 control stores of a retailer positioned in the middle that emphasizes customer-focused service as a way to gain strategic competitive advantage. The empirical methodology followed the two-stage technique in Banker and Natarajan (2008). First, using Data Envelopment Analysis (DEA), we computed the relative productivity of retail outlets in using their labor and capital resources (represented by total selling hours, store size, and average inventory) at generating (deflated) store sales. We then regressed the logarithm of DEA productivity scores on contextual variables to consistently estimate their impact on productivity and evaluate their statistical significance.

In agreement with our agency model, our empirical results indicate that supervisory monitoring and monetary incentives have a positive impact on retail sales productivity. Furthermore, we found that the marginal impact of monitoring is lower when the monetary incentives scheme is implemented (or, equivalently, that the value of monetary incentives decreases the higher is level of monitoring); and, as a result, that the level of monitoring is lower when monetary incentives are provided.
APPENDIX

Proof of Lemma 1:

The agent’s problem is to maximize with respect routine and creative effort the certainty equivalent

\[ CE[w_0(y(e_1, e_2)) ] = \alpha + \alpha_y (e_1 + e_2) - R\alpha_y^2\sigma_y^2(m)/2 - c_1e_1^2/2 - c_2e_2^2/2. \]

This yields the effort policies \( e_i = \alpha_y / c_i \) for \( i = 1, 2 \). The second-order conditions are satisfied because the effort cost function is convex. At the optimum, the individual rationality constraint binds, such that the expected compensation of the agent is

\[ \alpha + \alpha_y (e_1 + e_2) = r + R\alpha_y^2\sigma_y^2(m)/2 + c_1e_1^2/2 + c_2e_2^2/2. \]

Applying this expression to the objective of the principal, the principal’s problem becomes

\[ \max_{(\alpha_y)} e_1 + \lambda e_2 - r - R\sigma_y^2(m)\alpha_y^2/2 - c_1e_1^2/2 - c_2e_2^2/2. \]

Applying the effort policies and solving for the first-order condition, we obtain the expressions in the main text via substitution.

Proof of Proposition 1:

Taking the derivatives with respect to \( m \) of the expected outcome

\[ X_0(m) = \frac{(c_2 + \lambda c_1)^2/(c_1c_2)}{c_2 + c_1 + c_1c_2 R\sigma_y^2(m)} \] and expected profits \( \Pi_0(m) = \frac{(c_2 + \lambda c_1)^2/(2c_1c_2)}{c_2 + c_1 + c_1c_2 R\sigma_y^2(m)} - r \), we find that \( dX_0(m) / dm > 0 \) and \( d\Pi_0(m) / dm > 0 \) since \( d\sigma_y^2(m) / dm < 0 \).

Proof of Lemma 2:

The agent’s problem is to maximize with respect routine and creative effort the certainty equivalent

\[ CE[w_1(y(e_1, e_2), x(e_1, e_2))] = \beta + \beta_y (e_1 + e_2) + \beta_x(e_1 + \lambda e_2) - R[\beta_y^2\sigma_y^2(m) + \beta_x^2\sigma_x^2 + 2\beta_x\beta_y Cov]/2 - c_1e_1^2/2 - c_2e_2^2/2. \]

This yields the effort policies \( e_1 = (\beta_y + \beta_x)/c_1 \) and \( e_2 = (\beta_y + \lambda \beta_x)/c_2 \). The second-order conditions are satisfied because the effort cost function is convex. At the optimum, the individual rationality constraint binds, such that the expected compensation of the agent is
\[ \beta + \beta_y (e_1 + e_2) + \beta_x (e_1 + \lambda e_2) = r + R[\beta_y^2 \sigma^2_y (m) + \beta_x^2 \sigma^2_x + 2 \beta_x \beta_y \text{Cov}] / 2 + c_1 e_1^2 / 2 + c_2 e_2^2 / 2. \]

Applying this expression to the objective of the principal, the principal’s problem becomes
\[ \max_{(\beta_x, \beta_y)} e_1 + \lambda e_2 - R[\beta_y^2 \sigma^2_y (m) + \beta_x^2 \sigma^2_x + 2 \beta_x \beta_y \text{Cov}] / 2 - c_1 e_1^2 / 2 - c_2 e_2^2 / 2. \]

Applying the effort policies and solving for the first-order conditions, we obtain the expressions in the main text via substitution.

**Proof of Proposition 2:**

First, we simplify the expressions by defining \( A = c_2 + \lambda c_1 \), \( B = c_2 + \lambda^2 c_1 \), \( C = c_2 + c_1 \), \( D = c_1 c_2 R \text{Cov} \), \( E = c_1 c_2 R \sigma^2_x \), \( F = c_1 c_2 R \sigma^2_y (m) \), and \( V = BC - A^2 = c_1 c_2 (1 - \lambda)^2 \geq 0 \). Then we have that \( \beta_x = \frac{A(D + A) - B(F + C)}{(D + A)^2 - (E + B)(F + C)} \) and \( \beta_y = \frac{B(D + A) - A(E + B)}{(D + A)^2 - (E + B)(F + C)} \). Hence, we can express
\[ X_1 = \frac{1}{c_1 c_2} \left[ \frac{2ABD - A^2 E + A^2 B - B^2 C - B^2 F}{(D + A)^2 - (E + B)(F + C)} \right]. \]

We have that
\[ \frac{dX_1}{dm} = \frac{dX_1}{dF} \frac{dF}{dm} = R \frac{d}{dF} \left[ \frac{2ABD - A^2 E + A^2 B - B^2 C - B^2 F}{(D + A)^2 - (E + B)(F + C)} \right] \frac{d\sigma^2_y (m)}{dm} \]
\[ = \frac{-R(BD - AE)^2}{(D + A)^2 - (E + B)(F + C)} = \frac{d\sigma^2_y (m)}{dm} \cdot \frac{d\sigma^2_y (m)}{dm} \]

Because \( d\sigma^2_y (m) / dm < 0 \), we find that \( dX_1 (m) / dm > 0 \). Since \( \Pi_1 (m) = X_1 (m) / 2 - r \), this implies that \( d\Pi_1 (m) / dm > 0 \).

**Proof of Proposition 3:**

Using the same definitions as in the proof of Proposition 2, we have that the sign of \( X_1 - X_0 \) is determined by:
\[
X_t - X_0 = \frac{1}{c_1 c_2} \left[ \frac{2 ABD - A^2 E + A^2 B - B^2 C - B^2 F}{(D + A)^2 - (E + B)(F + C)} \right] - \frac{1}{c_1 c_2} \left( \frac{A^2}{C + F} \right).
\]

The numerator is negative. Define the term that determines the sign of the denominator as \( K = (D + A)^2 - (E + B)(F + C) \). We will show that \( K \leq 0 \). We have that
\[
K = 2AD + D^2 - EF - EC - V \left( \frac{A^2 + V}{C} \right) F,
\]
\[
= -\left( \sqrt{\frac{F}{C}} A - \sqrt{\frac{C}{F}} D \right)^2 - \left( \frac{FV}{C} + V \right) - \left( EF - D^2 \right) \left( \frac{F + C}{F} \right).
\]

Now, we have that
\[
(\text{EF} - D^2) = \left( c_1 c_2 R \sigma^2_y \right) \left( c_1 c_2 R \sigma^2_y(m) \right) = \left( c_1 c_2 R \right)^2 \left[ \sigma^2_y(m) - \text{Cov} \right].
\]

Because correlation cannot exceed one, we have \( \left[ \sigma^2_y(m) - \text{Cov} \right] \geq 0 \), such that \( (\text{EF} - D^2) \geq 0 \), which implies that \( K \leq 0 \). It follows that \( X_t - X_0 \geq 0 \). Since \( \Pi_1(m) = X_1(m) / 2 - r \) and \( \Pi_0(m) = X_0(m) / 2 - r \), the same holds for expected profits.

**Proof of Proposition 4:**

Using the same definitions as in the proof of Proposition 2, we take derivative of \( X_t - X_0 \) with respect to \( m \) to find that
\[
\frac{d}{dm} \left( X_t - X_0 \right) = \left( \frac{1}{c_1 c_2} \right) \frac{d}{dF} \left[ \frac{-[A(A + D) - B(C + F)]^2}{(D + A)^2 - (E + B)(F + C)} \right] \frac{dF}{dm}.
\]

Previously, we showed that \( \frac{dF}{dm} = \frac{1}{c_1 c_2 R \sigma^2_y(m)} = c_1 c_2 R \frac{d \sigma^2_y(m)}{dm} < 0 \). Thus, we only need to determine the sign of \( \frac{d}{dF} \left[ \frac{-[A(A + D) - B(C + F)]^2}{(D + A)^2 - (E + B)(F + C)} \right] : \)
\[
\frac{d}{dF} \left[ \frac{-[A(A + D) - B(C + F)]^2}{(D + A)^2 - (E + B)(F + C)} \right] = \frac{\frac{d}{dF} - [A(A + D) - B(C + F)]^2}{(D + A)^2 - (E + B)(F + C)} \frac{[A(A + D) - B(C + F)] - 2AEF + 2B(C + F)}{[D + A F + C]^2}.
\]

32
Therefore, \( \frac{d(X_f - X_0)}{dm} \leq 0 \) if and only if
\[
(A + D)[A(A + D) - B(C + F)] \left[ (A + D)[A(A + D) - B(C + F)] - 2A(E + B)(F + C) \right] \geq 0.
\]
Denoting \( M = (A + D)[A(A + D) - B(C + F)] \), we find that the inequality holds if \( M \leq 0 \).

The condition \( M \leq 0 \) holds if and only if \( A(A + D) - B(C + F) \leq 0 \), which may be expressed in terms of \( \lambda \) as follows:
\[
-c_1(c_1 + F)\lambda^2 + (2c_1c_2 + c_1D)\lambda_2 + c_2D - c_2(c_1 + F) \leq 0,
\]
where \( F \) and \( D \) do not depend on \( \lambda \). The roots of this quadratic function are
\[
\lambda_1 = \frac{(2c_1c_2 + c_1D) - \sqrt{(2c_1c_2 + c_1D)^2 + 4c_1(c_2 + F)[c_2D - c_2(c_1 + F)]}}{2c_1(c_2 + F)};
\]
\[
\lambda_2 = \frac{(2c_1c_2 + c_1D) + \sqrt{(2c_1c_2 + c_1D)^2 + 4c_1(c_2 + F)[c_2D - c_2(c_1 + F)]}}{2c_1(c_2 + F)}.
\]
Note that the first root is negative, but \( \lambda > 0 \). We infer that \( \frac{d(X_f - X_0)}{dm} \leq 0 \) if \( \lambda \geq \lambda_2 \), which is the condition in the main text. Since \( \Pi_f(m) = X_f(m)/2 - r \) and \( \Pi_0(m) = X_0(m)/2 - r \), the same holds for expected profits.

**Proof of Proposition 5:**

The optimal level of monitoring with outcome control is the solution to \( m_f = \arg\max_{[m]} \Pi_f(m) - D(m) \), while it is \( m_0 = \arg\max_{[m]} \Pi_0(m) - D(m) \) without outcome control. Since \( d(\Pi_f(m) - \Pi_0(m))/dm \leq 0 \) if \( \lambda \geq \lambda_2 \), it follows that \( m_0 \geq m_f \) if \( \lambda \geq \lambda_2 \).
REFERENCES


FIGURE 1: Input-Output Model with Contextual Variables

Inputs
- Total selling hours
- Store size
- Average inventory

Contextual Variables
- Monitoring and/or Monetary incentives scheme
- Average Sales index
- Holiday season
- Household income
- Age
- Blue collar workers
- Family size
- College education
- Population
- Rural area
- Competition

Output
- Store sales (Deflated)
### TABLE 1: Descriptive Statistics of Output, Inputs, and Contextual Variables (N=1,200)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Output</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SALES</td>
<td>$3.00M</td>
<td>$2.07M</td>
<td>$1.54M</td>
<td>$2.39M</td>
<td>$3.66M</td>
</tr>
<tr>
<td><strong>Inputs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOURS</td>
<td>39,173</td>
<td>21,133</td>
<td>22,645</td>
<td>34,461</td>
<td>48,977</td>
</tr>
<tr>
<td>SIZE</td>
<td>215.00</td>
<td>94.60</td>
<td>123.00</td>
<td>199.00</td>
<td>282.00</td>
</tr>
<tr>
<td>INVENTORY</td>
<td>$12.23M</td>
<td>$6.23M</td>
<td>$7.62M</td>
<td>$10.58M</td>
<td>$14.68M</td>
</tr>
<tr>
<td><strong>Contextual Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MONITORING</td>
<td>0.0873</td>
<td>0.0225</td>
<td>0.0712</td>
<td>0.0852</td>
<td>0.1009</td>
</tr>
<tr>
<td>SALESINDEX</td>
<td>$3.05M</td>
<td>$1.13M</td>
<td>$2.27M</td>
<td>$3.01M</td>
<td>$3.42M</td>
</tr>
<tr>
<td>INCOME</td>
<td>$37,944</td>
<td>$7,647</td>
<td>$31,766</td>
<td>$36,119</td>
<td>$43,922</td>
</tr>
<tr>
<td>POPLNAGE</td>
<td>32.59</td>
<td>2.15</td>
<td>31.50</td>
<td>32.35</td>
<td>33.85</td>
</tr>
<tr>
<td>BLUECOLLAR</td>
<td>42.59%</td>
<td>6.73%</td>
<td>38.05%</td>
<td>43.15%</td>
<td>48.25%</td>
</tr>
<tr>
<td>FAMILYSIZE</td>
<td>2.59</td>
<td>0.17</td>
<td>2.46</td>
<td>2.53</td>
<td>2.73</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>38.61%</td>
<td>8.19%</td>
<td>32.05%</td>
<td>39.10%</td>
<td>44.35%</td>
</tr>
<tr>
<td>POPULATION</td>
<td>353,882</td>
<td>172,424</td>
<td>221,029</td>
<td>384,586</td>
<td>496,624</td>
</tr>
<tr>
<td>RURAL</td>
<td>0.2500</td>
<td>0.4331</td>
<td>0</td>
<td>0</td>
<td>0.5000</td>
</tr>
<tr>
<td>COMPETITION</td>
<td>33.85</td>
<td>14.32</td>
<td>24.50</td>
<td>35.00</td>
<td>38.00</td>
</tr>
</tbody>
</table>

Variable definitions:
SALES = monthly deflated store sales, expressed in million dollars ($M);
HOURS = the number of monthly selling hours in a store;
SIZE = the size of a store in square feet;
INVENTORY = the deflated dollar value of the average inventory of merchandise, expressed in million dollars ($M);
MONITORING = the ratio of the number of managerial hours to the number of selling hours in a store;
SALESINDEX = monthly deflated average sales for 20 stores;
INCOME = median household income of the county in which a store is located;
POPLNAGE = median age of the population of the county in which a store is located;
BLUECOLLAR = the percentage of blue collar workers of the county in which a store is located;
FAMILYSIZE = median family size of the county in which a store is located;
COLLEGE = the percentage of the population with a college education for the county in which a store is located;
POPULATION = population size in the county in which a store is located;
RURAL = 1 if a store is located in a rural area, and 0 otherwise;
COMPETITION = the competition intensity of a store’s market.
<table>
<thead>
<tr>
<th></th>
<th>Stores with monetary incentives scheme</th>
<th>Stores without monetary incentives scheme</th>
<th>p-value for t-test of difference of mean=0</th>
<th>p-value for sign test of difference of median=0</th>
<th>p-value for sign rank test of difference of median=0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Output</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SALES</td>
<td>$3.05M ($2.78M)</td>
<td>$3.02M ($2.84M)</td>
<td>0.826</td>
<td>1.000</td>
<td>0.922</td>
</tr>
<tr>
<td><strong>Inputs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOURS</td>
<td>37,932 (35,123)</td>
<td>35,495 (33,467)</td>
<td>0.276</td>
<td>0.344</td>
<td>0.193</td>
</tr>
<tr>
<td>SIZE</td>
<td>203.80 (183.50)</td>
<td>226.20 (227.50)</td>
<td>0.133</td>
<td>0.109</td>
<td>0.131</td>
</tr>
<tr>
<td>INVENTORY</td>
<td>$12.62M ($11.84M)</td>
<td>$12.31M ($10.76M)</td>
<td>0.680</td>
<td>0.754</td>
<td>0.625</td>
</tr>
<tr>
<td><strong>Contextual Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MONITORING</td>
<td>0.0928 (0.0900)</td>
<td>0.0943 (0.0882)</td>
<td>0.576</td>
<td>0.754</td>
<td>0.492</td>
</tr>
<tr>
<td>INCOME</td>
<td>$39,590 ($39,126)</td>
<td>$37,662 ($35,941)</td>
<td>0.576</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>POPLNAGE</td>
<td>32.39 (32.30)</td>
<td>32.78 (32.95)</td>
<td>0.563</td>
<td>0.754</td>
<td>0.322</td>
</tr>
<tr>
<td>BLUECOLLAR</td>
<td>41.05% (41.00%)</td>
<td>44.12% (44.35%)</td>
<td>0.304</td>
<td>0.344</td>
<td>0.432</td>
</tr>
<tr>
<td>FAMILYSIZE</td>
<td>2.60 (2.51)</td>
<td>2.59 (2.58)</td>
<td>0.919</td>
<td>1.000</td>
<td>0.988</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>40.31% (40.05%)</td>
<td>36.90% (38.00%)</td>
<td>0.367</td>
<td>0.344</td>
<td>0.432</td>
</tr>
<tr>
<td>POPULATION</td>
<td>335,704 (391,331)</td>
<td>372,059 (346,500)</td>
<td>0.649</td>
<td>0.754</td>
<td>0.770</td>
</tr>
<tr>
<td>RURAL</td>
<td>0.4000 (0)</td>
<td>0.1000 (0)</td>
<td>0.193</td>
<td>0.375</td>
<td>0.375</td>
</tr>
<tr>
<td>COMPETITION</td>
<td>31.50 (35.00)</td>
<td>36.20 (36.50)</td>
<td>0.532</td>
<td>0.727</td>
<td>0.359</td>
</tr>
</tbody>
</table>
| TABLE 3: Matched Pair Tests for the Post-Implementation Period  
(means, with medians in parentheses below) | Stores with monetary incentives scheme | Stores without monetary incentives scheme | p-value for t-test of difference of mean=0 | p-value for sign test of difference of median=0 | p-value for sign rank test of difference of median=0 |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Output</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SALES</td>
<td>$3.02M ($2.50M)</td>
<td>$2.85M ($2.62M)</td>
<td>0.266</td>
<td>1.000</td>
<td>0.846</td>
</tr>
<tr>
<td><strong>Inputs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOURS</td>
<td>43,222 (38,635)</td>
<td>38,516 (35,871)</td>
<td>0.061</td>
<td>0.022</td>
<td>0.004</td>
</tr>
<tr>
<td>SIZE</td>
<td>203.80 (183.50)</td>
<td>226.20 (227.50)</td>
<td>0.133</td>
<td>0.109</td>
<td>0.131</td>
</tr>
<tr>
<td>INVENTORY</td>
<td>$12.38M ($11.25M)</td>
<td>$11.49M ($9.94M)</td>
<td>0.211</td>
<td>0.754</td>
<td>0.232</td>
</tr>
<tr>
<td><strong>Contextual Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MONITORING</td>
<td>0.0699 (0.0686)</td>
<td>0.0905 (0.0893)</td>
<td>0.000</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>INCOME</td>
<td>$38,579 ($38,369)</td>
<td>$35,370 ($33,477)</td>
<td>0.399</td>
<td>1.000</td>
<td>0.432</td>
</tr>
<tr>
<td>POPLNAGE</td>
<td>32.39 (32.30)</td>
<td>32.78 (32.95)</td>
<td>0.563</td>
<td>0.754</td>
<td>0.322</td>
</tr>
<tr>
<td>BLUECOLLAR</td>
<td>41.05% (41.00%)</td>
<td>44.12% (44.35%)</td>
<td>0.304</td>
<td>0.344</td>
<td>0.432</td>
</tr>
<tr>
<td>FAMILYSIZE</td>
<td>2.60 (2.51)</td>
<td>2.59 (2.58)</td>
<td>0.919</td>
<td>1.000</td>
<td>0.945</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>40.31% (40.05%)</td>
<td>36.90% (38.00%)</td>
<td>0.367</td>
<td>0.344</td>
<td>0.432</td>
</tr>
<tr>
<td>POPULATION</td>
<td>335,704 (391,331)</td>
<td>372,059 (346,500)</td>
<td>0.649</td>
<td>0.754</td>
<td>0.770</td>
</tr>
<tr>
<td>RURAL</td>
<td>0.4000 (0)</td>
<td>0.1000 (0)</td>
<td>0.193</td>
<td>0.375</td>
<td>0.375</td>
</tr>
<tr>
<td>COMPETITION</td>
<td>31.50 (35.00)</td>
<td>36.20 (36.50)</td>
<td>0.532</td>
<td>0.727</td>
<td>0.359</td>
</tr>
</tbody>
</table>
### TABLE 4: DEA efficiency scores

<table>
<thead>
<tr>
<th></th>
<th>Pooled Data (N=1,200)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0.2036</td>
</tr>
<tr>
<td>Maximum</td>
<td>1.0000</td>
</tr>
<tr>
<td>Mean</td>
<td>0.5245</td>
</tr>
<tr>
<td>First quartile</td>
<td>0.4155</td>
</tr>
<tr>
<td>Median</td>
<td>0.4936</td>
</tr>
<tr>
<td>Third quartile</td>
<td>0.5893</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.1664</td>
</tr>
<tr>
<td>Number of observations on efficient frontier</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>( \ln \hat{\theta} )</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>( \ln \hat{\theta} )</td>
<td>-0.0021</td>
</tr>
<tr>
<td>MONITOR</td>
<td>0.0402</td>
</tr>
<tr>
<td>PP</td>
<td>-0.1791</td>
</tr>
<tr>
<td>PP_MONITOR</td>
<td>-0.2204</td>
</tr>
<tr>
<td>SINDEX</td>
<td>0.6675</td>
</tr>
<tr>
<td>SEASON</td>
<td>0.4951</td>
</tr>
<tr>
<td>INCOME</td>
<td>-0.0794</td>
</tr>
<tr>
<td>AGE</td>
<td>-0.1451</td>
</tr>
<tr>
<td>BLUE</td>
<td>-0.0153</td>
</tr>
<tr>
<td>FAMSIZE</td>
<td>-0.1561</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>0.1246</td>
</tr>
<tr>
<td>POPUL</td>
<td>-0.1800</td>
</tr>
<tr>
<td>RURAL</td>
<td>-0.0615</td>
</tr>
<tr>
<td>COMPET</td>
<td>-0.1824</td>
</tr>
</tbody>
</table>
TABLE 6
Results of Regressing the Logarithm of Productivity Estimates on Contextual Variables

\[
\ln \hat{\theta}_{jt} = \beta_0 + \beta_1 \times \text{MONITORING}_{jt} + \beta_2 \times \text{MONINCENT}_{jt} \\
+ \beta_3 \times \text{MONINCENT}_{jt} \times \text{MONITORING}_{jt} + \beta_4 \times \text{AVGSAL}_t + \beta_5 \times \text{SEASON}_t \\
+ \beta_6 \times \text{INCOME}_j + \beta_7 \times \text{POPLNAGE}_j + \beta_8 \times \text{BLUECOLLAR}_j + \beta_9 \times \text{FAMILYSIZE}_j \\
+ \beta_{10} \times \text{COLLEGE}_j + \beta_{11} \times \text{POPULATION}_j + \beta_{12} \times \text{RURAL}_j + \beta_{13} \times \text{COMPETITION}_j + \epsilon_{jt}
\]

Panel A: Parameter Estimates (p-value in parentheses)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Expected sign</th>
<th>Parameter Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>-0.3484***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0001)</td>
</tr>
<tr>
<td>MONITORING</td>
<td>+</td>
<td>2.4052***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>MONINCENT</td>
<td>+</td>
<td>0.1850***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0032)</td>
</tr>
<tr>
<td>MONINCENT*MONITORING</td>
<td>-</td>
<td>-3.5879***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>AVGSALES (in millions)</td>
<td>+</td>
<td>0.1616***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>SEASON</td>
<td>+</td>
<td>0.1868***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>INCOME (in millions)</td>
<td>+</td>
<td>13.5700***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>POPLNAGE</td>
<td>+</td>
<td>-0.0147***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0011)</td>
</tr>
<tr>
<td>BLUECOLLAR</td>
<td>+</td>
<td>0.0135***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>FAMILYSIZE</td>
<td>-</td>
<td>-0.5378***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0000)</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>+</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.1871)</td>
</tr>
<tr>
<td>POPULATION (in millions)</td>
<td>-</td>
<td>-0.0192</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.4202)</td>
</tr>
<tr>
<td>RURAL</td>
<td>+</td>
<td>0.0618***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0089)</td>
</tr>
<tr>
<td>COMPETITION</td>
<td>-</td>
<td>-0.0029***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0007)</td>
</tr>
<tr>
<td>p(model)</td>
<td></td>
<td>0.0001</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td></td>
<td>0.6845</td>
</tr>
</tbody>
</table>

*, ** and *** indicate statistical significance at 10%, 5% and 1% levels respectively for one-sided hypothesis tests.
### Panel B: Hypothesis Tests

<table>
<thead>
<tr>
<th>( H1 ):</th>
<th>( p(\alpha_1 &gt; 0) )</th>
<th>0.0285.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H2 ):</td>
<td>( p(\alpha_3 &lt; 0) )</td>
<td>0.0304</td>
</tr>
<tr>
<td></td>
<td>( p(\alpha_1 + \alpha_3 &lt; 0) )</td>
<td>0.1934</td>
</tr>
<tr>
<td>( H3 ):</td>
<td>( p(\alpha_2 + \alpha_3 \cdot \text{mean of } \Delta \text{MONITORING} &gt; 0) )</td>
<td>0.2879</td>
</tr>
<tr>
<td></td>
<td>( p(\alpha_2 + \alpha_3 \cdot 1^{\text{st}} \text{ decile of } \Delta \text{MONITORING} &gt; 0) )</td>
<td>0.0360</td>
</tr>
<tr>
<td></td>
<td>( p(\alpha_2 + \alpha_3 \cdot \text{lower quartile of } \Delta \text{MONITORING} &gt; 0) )</td>
<td>0.0815</td>
</tr>
<tr>
<td></td>
<td>( p(\alpha_2 + \alpha_3 \cdot \text{median of } \Delta \text{MONITORING} &gt; 0) )</td>
<td>0.2195</td>
</tr>
<tr>
<td></td>
<td>( p(\alpha_2 + \alpha_3 \cdot \text{upper quartile of } \Delta \text{MONITORING} &lt; 0) )</td>
<td>0.4732</td>
</tr>
<tr>
<td></td>
<td>( p(\alpha_2 + \alpha_3 \cdot 9^{\text{th}} \text{ decile of } \Delta \text{MONITORING} &lt; 0) )</td>
<td>0.1186</td>
</tr>
</tbody>
</table>

**Variables definitions:**

\( \ln \hat{\theta}_{jt} \) = the logarithm of DEA productivity score for store \( j \) in month \( t \),

\( \text{MONITORING}_{jt} \) = the ratio of the number of managerial hours to the number of total selling hours in store \( j \) in month \( t \);

\( \text{MONINCENT}_{jt} \) = 1 if store \( j \) was under the monetary incentives scheme in month \( t \), and 0 otherwise;

\( \text{AVGSALES}_t \) = monthly deflated average sales for all 20 stores;

\( \text{SEASON}_t \) = 1 if month \( t \) is October, November, or December, and 0 otherwise;

\( \text{INCOME}_j \) = median household income of the county in which store \( j \) is located, expressed in million dollars;

\( \text{POPLNAGE}_j \) = median age of the population for the county in which store \( j \) is located;

\( \text{BULECOLLAR}_j \) = the percentage of blue collar workers for the county in which store \( j \) is located;

\( \text{FAMILYSIZE}_j \) = median family size for the county in which store \( j \) is located;

\( \text{COLLEGE}_j \) = the percentage of the population with a college education for the county in which store \( j \) is located;

\( \text{POPULATION}_j \) = the population size in the county in which store \( j \) is located, expressed in millions;

\( \text{RURAL}_j \) = 1 if store \( j \) is located in a rural area, and 0 otherwise;

\( \text{COMPETITION}_j \) = the competition intensity of store \( j \)’s market.
Table 7
Results of Regressing the Change in the Logarithm of
Productivity Estimates on the Change in Contextual Variables

\[ \Delta \ln \theta_{jt} = \alpha_0 + \alpha_1 \Delta \text{MONITORING}_{jt} + \alpha_2 \text{MONINCENT}_{jt} \]

\[ + \alpha_3 \text{MONINCENT}_{jt} \Delta \text{MONITORING}_{jt} + \alpha_4 \Delta \text{AVGEFFICIENCY}_{jt} \]

\[ + \alpha_5 \Delta \text{SALESGROWTH}_{jt} + \alpha_6 \Delta \text{AVGSALES}_{jt} + \varepsilon_{jt} \]

<table>
<thead>
<tr>
<th>Parameter Estimates (p-value in parentheses)</th>
<th>Parameter</th>
<th>Expected sign</th>
<th>Parameter Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>( \alpha_0 )</td>
<td></td>
<td>-0.0005 (0.9241)</td>
</tr>
<tr>
<td>( \Delta \text{MONITORING} )</td>
<td>( \alpha_1 )</td>
<td>+</td>
<td>0.6996** (0.0285)</td>
</tr>
<tr>
<td>( \text{MONINCENT} )</td>
<td>( \alpha_2 )</td>
<td>+</td>
<td>0.0033 (0.3109)</td>
</tr>
<tr>
<td>( \text{MONINCENT} \Delta \text{MONITORING} )</td>
<td>( \alpha_3 )</td>
<td>+</td>
<td>-1.2053** (0.0304)</td>
</tr>
<tr>
<td>( \Delta \text{AVGEFFICIENCY} )</td>
<td>( \alpha_4 )</td>
<td>+</td>
<td>1.0044*** (0.0000)</td>
</tr>
<tr>
<td>( \Delta \text{SALESGROWTH} )</td>
<td>( \alpha_5 )</td>
<td>+</td>
<td>-0.0615 (0.3929)</td>
</tr>
<tr>
<td>( \Delta \text{AVGSALES} ) (in millions)</td>
<td>( \alpha_6 )</td>
<td>+</td>
<td>0.0016 (0.4094)</td>
</tr>
<tr>
<td>( \text{p(model)} )</td>
<td></td>
<td></td>
<td>0.0001</td>
</tr>
<tr>
<td>Adjusted ( R^2 )</td>
<td></td>
<td></td>
<td>0.8099</td>
</tr>
</tbody>
</table>

*, ** and *** indicate statistical significance at 10%, 5% and 1% levels respectively for one-sided hypothesis tests.

Variables definitions:
\( \Delta \ln \theta_{jt} \) = change in the logarithm of productivity measure for store \( j \) from month \( t \) to \( t-1 \),
\( \Delta \text{MONITORING}_{jt} \) = change in the ratio of the number of managerial hours to the number of total selling hours in a store \( j \) from month \( t \) to \( t-1 \),
\( \text{MONINCENT}_{jt} = 1 \) if store \( j \) was under the monetary incentives scheme in month \( t \), 0 otherwise,
\( \Delta \text{AVGEFFICIENCY}_{t} \) = change in monthly average of the logarithm of productivity measures for 20 stores from month \( t \) to \( t-1 \),
\( \Delta \text{SALESGROWTH}_{jt} \) = change in sales growth for store \( j \) from month \( t \) to \( t-1 \),
\( \Delta \text{AVGSALES} \) = change in average monthly sales for 20 stores from month \( t \) to \( t-1 \).