SUNIL KISHORE, RAGHUNATH SINGH RAO, OM NARASIMHAN, and GEORGE JOHN*

Quota-based bonuses and commissions are the two most common incentive compensation plans. The authors uncover differential effects of these plans from a natural field-based experiment featuring 14,000 monthly observations over three years from 458 sales territories of a pharmaceutical firm that switched from a bonus plan to an equivalent commission plan. The intervention led to significant sales productivity improvement; this effect was heterogeneous across ability deciles, with much larger increases occurring at lower ability deciles. The authors find significant differences across these plans on (1) effort against nonincentivized tasks and (2) output fluctuations induced through "timing games." At this firm, the bonus plan was strictly inferior to the implemented commission plan with respect to short-term revenues and timing games. In contrast, the commission plan induced greater neglect of nonincentivized tasks (tasks not directly affecting observable output). To organize their findings, the authors build a simple theoretical model in the personnel economics tradition. The novel result that multitasking concerns are reduced under bonus plans when the quota has been met provides a nuanced rationale for the widespread existence of lump-sum bonus plans.

Keywords: sales compensation, field experiments, timing games, multitasking

Bonuses Versus Commissions: A Field Study

Firms have long recognized the importance of designing appropriate incentive plans to induce their salespeople to exert effort. Extensive theoretical literature in marketing and economics that is centered on principal–agent models has involved attempts to understand how different incentive compensation schemes affect effort choices, with the aim of characterizing optimal incentive contracts (for a review of studies in economics, see Prendergast 1999; for studies in marketing, see Albers and Mantrala 2008). Although the theoretical work yields several important insights into the design of incentive plans, there is surprisingly little empirical work on the efficacy of various incentive schemes firms actually use in practice.

In the real world, most incentive plans are invariably nonlinear, and the variable component begins when salespeople meet a threshold that is set above the minimum output level, commonly known as a "quota" in the sales management literature. Joseph and Kalwani's (1998) survey of sales compensation practices at *Fortune* 500 companies indicates that 95% of the firms used plans that featured a quota. Quotas have been rationalized as devices to account for salesperson territory heterogeneity (e.g., Zoltners, Prabhakant, and Lorimer 2008). In addition, psychological goal theory (for a comprehensive review, see Latham and Locke 1991) suggests that quotas provide salespeople with challenging objectives when pursued and with psychological rewards when attained (Jain 2009).

Two forms of variable pay are widely used in conjunction with quotas: (1) lump sums (bonuses) paid upon reaching quota and (2) per-unit payouts (commissions) paid on sales beyond the quota (see Figure 1). For ease of exposition, we

*Sunil Kishore is an associate at McKinsey & Company (e-mail: kisho013@umn.edu). Raghunath Singh Rao is Assistant Professor of Marketing, McCombs School of Business, University of Texas at Austin (e-mail: raghunath.rao@mccombs.utexas.edu). Om Narasimhan is Professor, Department of Management, London School of Economics (e-mail: o.narasimhan@lse.ac.uk). George John is Pillsbury-Gerot Chair in Marketing, Carlson School of Management, University of Minnesota, and Distinguished Visiting Professor, King Abdul-Aziz University, Jeddah, Saudi Arabia (e-mail: johnx001@umn.edu). The authors acknowledge helpful comments from seminar participants at the University of Texas, the University of Arizona, Stanford University, the 2011 China–India Consumer Insights Conference in New Delhi, and the 2012 Marketing Science Conference in Boston. They are grateful to the executives of the focal firm (who wish to remain anonymous) for providing the data and for their helpful discussions. This article is based on the first author's doctoral dissertation at the University of Minnesota. The first two authors contributed equally and are listed alphabetically. Jean-Pierre Dubé served as associate editor for this article.

ISSN: 0022-2437 (print), 1547-7193 (electronic)

DOI: 10.1509/jmr.11.0485

hereinafter refer to these as "bonus" and "commission" plans, respectively, with the understanding that both involve (nonnegative) quotas. Intuitively, a bonus plan provides a strong incentive for salespeople to reach the quota, but it also encourages them to minimize their efforts thereafter, whereas a commission plan provides an incentive to keep working hard even after reaching the quota. Oyer (1998) and Jensen (2003) formalize these insights with models showing that bonuses tempt salespeople to manipulate the timing of orders (e.g., delayed selling, forward selling). As such, it seems that bonus plans are perhaps strictly inferior to linear commissions if output is the sole criterion used to judge a plan's efficacy (for an exception, see Oyer 2000). Previous theoretical conjectures about the relative merits of these plans notwithstanding, the empirical literature is sparse and inconclusive on the subject. The empirical issue that has received some systematic attention in this area is that of deliberate postponement or advancement of sales by salespeople (referred to as "timing games") (Misra and Nair 2011; Oyer 1998; Steenburgh 2008).

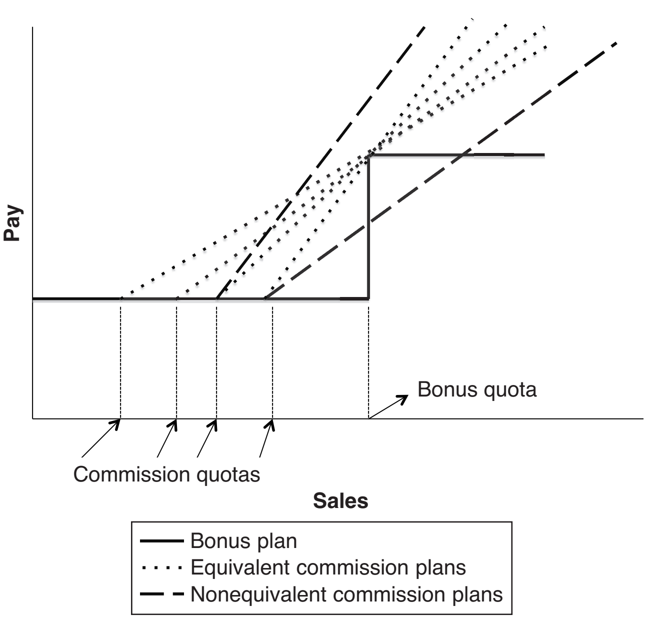
This empirical literature has largely ignored the "multitasking" distortion associated with high-powered incentives (Holmstrom and Milgrom 1991). We previously stated that commission plans seem to provide stronger incentives than bonus plans on measurable tasks used in the calculation of incentive pay. As such, if salespeople are assigned a portfolio of tasks—some of which are not explicitly included in the computation of incentive pay (either because of the unavailability of good outcome measures or because it is easy for agents to "shade" these measures)—commission plans should induce agents to neglect these other tasks to a greater degree than would bonus plans.

We use data from a large firm-level field experiment to study the impact on productivity as well as the impact on the focal firm's output fluctuations and multitasking concerns. The experiment involved switching the firm's sales compensation from bonuses to linear commissions. From a methodological viewpoint, researchers have increasingly used field experiments to study contracting (e.g., Bandiera, Barankay, and Rasul 2005; Paarsch and Shearer 2000) because they alleviate endogeneity concerns. A few researchers have also recently tried to estimate the structural parameters of sales contracts from observational data (Chung, Steenburgh, and Sudhir 2010; Misra and Nair 2011). Neither of these studies contrasts bonuses with commissions, which is the objective of our article. More specifically, we designed a quasi-experiment to address the following questions:

1. What are the productivity consequences of bonus plans compared with commissions? We are interested in the overall effect as well as differences in impact across agents.
2. How do timing games manifest themselves under bonuses and commissions? Sales fluctuations induced by incentive plans can be inefficient for organizations (Jensen 2003). Our study sheds light on sales timing behavior.
3. Under which plan (bonuses or commissions) are multitasking concerns more severe? Although the neglect of nonincentivized tasks under high-powered incentives is well-known conceptually, researchers know little about them empirically across bonus and commission plans.

Figure 1

PLOTS OF INCENTIVE COMPENSATION PLANS



The following is a preview of our findings: (1) The commission plan increases sales force productivity by approximately 24% over the bonus plan. Specifically, the switch has heterogeneous impact: it leads to higher improvements at lower ability deciles. (2) Salespeople "push sales into the future" if they are unlikely to meet quota and "pull sales in from the future" if they are near quota. These patterns are more pronounced under the bonus plan. (3) After salespeople have met a quota, the bonus scheme is much better than commissions in terms of engaging salespeople to undertake tasks that are not directly compensated but are important to the firm.

An immediate issue that arises from single-firm field studies such as the one we propose here pertains to generalizability. Consider each of the three previous questions in turn. The generalizability of our productivity predictions (Question 1) turns on defining plan equivalence. Suppose the quota a firm sets is difficult to reach under bonuses. Now, if the firm were to switch to a commission plan in which the quota was set at a relatively low level of sales, it is intuitively evident that the commission plan would likely improve productivity, but this would not generalize to other commission (or bonus) plans. As such, we compare bonus and commission schemes that are "equivalent" (a term we define precisely subsequently). It is important to note that our comparison does not require either plan to be "optimal"; we only require data on two equivalent plans. The case for generalizability is even stronger for our timing games (Question 2) and multitasking distortion (Question 3) predictions. In neither case do we require plans to be either optimal or equivalent. For example, salespeople can distort the sale in the final month of a quarter by pulling in sales from the future or pushing out sales into the future, and we predict that after the quota is met, the former (latter) is more likely to occur under commissions (bonuses). This result is general and does not rely on the process that went into the firm's incentive design or whether the contracts are equivalent. Taken together, these theoretical elements enable generalizability even with data from a single-firm design. That said, we acknowledge that a single-firm experiment permits detailed data collection and controlled implementation, albeit at the cost of context generalizability.

Regarding the implementation of the study, we successfully urged a division of a large pharmaceutical firm in an emerging economy to change its extant bonus plan to a commission plan. As a result, we have a before–after quasi-experimental design; that is, the change in scheme was exogenous. We exploit this exogenous change to identify the unique effects of the two plans.

We organize the rest of the article as follows: In the next section, we describe our empirical context and follow with a stylized theoretical model inspired by our empirical setting. Then, we report empirical findings and conclude the article with a discussion of our findings' implications. Details of the theory model appear in the self-contained Web Appendix A. Web Appendix B provides a detailed discussion of data and identification issues and robustness checks. (For both Web Appendixes, see www.marketingpower.com/jmr_webappendix.)

EMPIRICAL SETTING

We obtained detailed territory-level data from a large pharmaceutical firm in an emerging market. This firm sells drugs in multiple therapeutic classes and employs a dedicated sales force for each class. We worked with the sales force for one of these therapeutic classes. Our data consist of observations about sales groups of varying sizes assigned to territories, with unique quarterly quotas/targets for each territory.[1] The firm was actively examining changes to its extant incentive pay plan for salespeople, which consisted of lump-sum bonuses paid when the territory sales reached the quota assigned at the beginning of that quarter. We worked with the firm to set up a commission scheme taking into consideration our theoretical framework.

Institutional Details

There are three sets of actors in this market: salespeople, pharmacies, and distributors. The salespeople who constitute a sales group are responsible for a specific territory. The number of salespeople in a sales group is determined by the market potential of the territory, which is roughly a function of the number of physicians and pharmacies operating in the territory. Each salesperson is assigned a specific list of physicians and pharmacies. The salesperson calls on doctors to encourage them to prescribe the firm's drugs (referred to as "doctor visits") and pharmacies to encourage them to buy and stock the firm's drugs ("pharmacy visits"). However, sales cannot be parsed unambiguously; when a customer purchases a drug from a pharmacy, the firm cannot trace it back to a particular physician (and thus credit a specific salesperson). Note that this is not an issue in single-salesperson territories, because all sales in the territory are credited to that salesperson. Each salesperson reports the number of doctor and pharmacy visits made daily in a call report system. Managers conduct random postreport checks by calling or visiting physicians and pharmacies mentioned in the salesperson's call report.

Pharmacies are typically single-owner establishments with a few employees. The owner (or an employee) is required to have a pharmacy certification. There are multiple pharmacies within each territory, and they are quite competitive with one another. They stock a limited volume of drugs and rely on frequent deliveries from distributors. Distributors are midsized trading firms that buy the drugs from the firm and are responsible for ensuring that drugs are delivered to the pharmacies. There are typically multiple distributors within each territory.

Incentive compensation is computed by crediting all sales in a territory equally. All salespeople in the territory get equal credit; thus, the incentive compensation is split equally. Salespeople in a territory routinely meet and exchange information on detailing and sales achievements. It is crucial at this point to emphasize that even with multiple salespeople working in a single territory, there is no team selling as such. Although salespeople meet and compare notes routinely, they do not make joint calls or otherwise complement one another. We expected concerns about free riding in multiperson territories (Hardin 1968) given the firm's equal credit allocation for sales, but the managers we interviewed told us that free riding was not a major problem. Our interviews with salespeople suggest a possible reason. The salespeople in a territory meet routinely for office conferences. Presumably, the peer monitoring that occurs in these meetings dampens free riding, as previous research has shown (e.g., Hamilton, Nickerson, and Owan 2003; Knez and Simester 2001); we examine this issue further when discussing our empirical results.

Doctor visits. Each salesperson visits the doctors assigned to him or her. These visits are a classic example of a "pull" marketing tactic. During a typical visit (which can last from 5 to 20 minutes), the salesperson highlights the effectiveness of the drugs, distributes technical literature, and provides free samples. Another way to conceptualize these visits is akin to *informative* selling, wherein provision of information is used as a tool to encourage doctors to prescribe the drugs. Although these visits undoubtedly build preference and loyalty for the company's products, they do not translate into immediate sales because of two crucial institutional features. First, physicians in the United States may legally write only the drug's active ingredient on the prescription. Typically, there are multiple equivalent brands as a result of limited patent protection.[2] Second, even when a physician prescribes a particular brand, a pharmacist may legally substitute a different brand as long as it is chemically equivalent. Despite these factors, the importance of the physician remains; patients are rarely assertive and generally defer to physicians. The reasons for this are partly cultural, in that physicians are high-status people, and partly institutional, in that direct-to-customer drug advertising is not permitted by law.

[1]To preserve the anonymity of the company that provided the personnel records, we have scaled all the time-series data by a single common positive number. All monetary values in the main body of the article and the tables are in the hundreds of thousands in local currency except the industry sales, which are in the tens of millions in local currency.

[2]In this market, the legal system allows only process patents and not product patents. As soon as a new drug is launched, several equivalent products manufactured using noninfringing processes are launched fairly quickly.

Pharmacists tend to stock a limited number of brands from the plethora of options available. Therefore, when a patient walks into a pharmacy, he or she is likely to buy from among the brands that are in stock at the pharmacy. In addition, although most of the instructions on the package are in English, many patients are not fluent in English, giving the pharmacist considerable influence.

If a physician has a high preference for a particular brand (as a result of salesperson-provided information, among other things) and conveys that to patients, patients might insist on getting that particular brand at the retail level. This also results in pharmacies being more likely to stock these brands. In short, although marketing efforts made at the physician level may not translate into immediate sales at the retail level, they do have important consequences for the firm. Given the importance of doctor visits, but absent a clear output measure, the firm enforces a minimum level of self-reported doctor visits per quarter and imposes a penalty if this level is not met.

Pharmacy visits. Pharmacy visits are a "push" marketing tool for this firm. A typical pharmacy visit is longer than a doctor visit (25–45 minutes). During such a visit, the salesperson and the pharmacy owner discuss sales, stock levels, future orders, credit, and other trade terms. Salespeople often also obtain competitive intelligence and pricing information at the pharmacy, because pharmacists interact with salespeople from multiple manufacturers. A pharmacy visit is less formal than a doctor visit; among other things, the timing and volume of a booked sale is often influenced by a salesperson's ties with a pharmacy owner. These efforts can be thought of as *persuasive* selling, because convincing a pharmacy to stock more of the firm's products results in booking a sale that is immediately credited to the salesperson and affects his or her compensation. Salespeople have some pricing authority within a band and can negotiate terms of trade, such as the credit period, billing cycle, samples, and more, which has important implications for the timing games we discuss subsequently.

Market Conditions

The prescription drug market has increased at double-digit rates in the past decade as access to health care in the focal emerging market has grown rapidly. Although primary demand is influenced by many factors, the medical representatives are the principal marketing resource deployed by drug firms. Our focal category of products treats a specific nonchronic ailment whose demand at the patient level is subject to seasonal variation. Patients do not stockpile these drugs because they are prescribed for a particular period (typically two to three weeks). Stockpiling by pharmacies is limited because prices are relatively stable and products have an expiration date. We do not observe wholesale prices, but our understanding is that they were fairly stable during the period of our analysis. Two additional notes about prices are important: (1) Firms do not seem to compete on price; prices across competing brands are relatively similar. (2) No firm is large enough to influence prices at the market level: within this category, there were at least 15 firms offering products, and market shares rarely exceeded 10% for any firm.

Given the aforementioned notes, it seems that a crucial driver of demand at the individual firm level is sales force management—both the intensity of selling efforts and the way salespeople engage in informative (i.e., doctor visits) and persuasive (i.e., pharmacy visits) aspects of the selling process. With regard to overall market conditions, the number of pharmacies salespeople visit in each territory remained stable over the observation period. Finally, even though the overall market was increasing, the size of sales groups within the period of our analysis remained largely stable (see Figure B1 in the Web Appendix at www.marketingpower.com/jmr_webappendix).

Experimental Design

The focal firm employed an incentive plan wherein each salesperson received a monthly salary and a quarterly lump-sum bonus upon achievement of quota. After discussing the pros and cons of different incentive plans, we persuaded the firm to shift to a commission plan in which each salesperson received a fixed monthly salary and earned commissions at a fixed rate on any sales exceeding commission quota. We devised this new plan following a two-step procedure.

First, we asked the firm to set quotas for their extant (bonus) plan in line with their existing quota-setting process (see details of quota setting process in Web Appendix B at www.marketingpower.com/jmr_webappendix) for the quarter at the beginning of the intervention. Next, using these quotas as a benchmark, we asked the firm to revise the quotas for a commission plan such that a sales group achieving the quota from the first step would earn the same amount of incentive pay under the new plan. This two-step procedure ensured "equivalence" of the old and new plans, a concept we define more precisely in the next section. The revised quotas under the commission plan were lower than the quotas under the bonus plan set in the first step (as is required by our theoretical setup). Parenthetically, there are many possible combinations of quota and commission rates that would generate an equivalent treatment plan (for illustration of several possible commission plans, both equivalent and non-equivalent, see Figure 1),[3] and we left it to the firm to create its preferred combination of quota and commission rate.[4]

In addition, we designed the experiment to minimize experimental artifacts. In particular, we were concerned about Hawthorne effects and differential attrition problems, which are common in field settings. To manage these issues, we (encouraged by the firm) eschewed a treatment–control groups design because such designs highlight the contrast between the treated and control groups. To manage attrition artifacts, we persuaded the firm to launch the new plan in

[3]We operationalized an "equivalent" scheme as follows: In the preintervention period, the quarterly lump-sum bonus paid in each territory was approximately 2.5% of each bonus sales quota. For the first quarter of the postregime period, the firm calculated the new bonus sales quota (Q_{iBN}). At this stage, we provided several arithmetic examples of equivalent commission schemes and quizzed the managers to ensure that they understood equivalence unambiguously. For the postregime period, the firm's managers first settled on a commission rate of 8.33% (for sales above the commission quota) for every territory. Then, to maintain equivalence with the preregime period, they set the new quotas as $Q_{iBN} \times (2.5\%) = (8.33\%)(Q_{iBN} - Q_{icN})$, which yields $Q_{icN} = .7(Q_{icB})$. In subsequent quarters, these quotas were updated following established procedures (see Web Appendix B at www.marketingpower.com/jmr_webappendix).

[4]The final choice of the commission scheme from the many possibilities was dictated by the existing quota-setting process within the firm and the expected incentive payout, among other things.

exactly the same way the firm typically does when making changes to compensation. As we document subsequently, our efforts seemed to be successful, with no measurable changes in quit rates.[5] The availability of repeated observations of pre- and postexperiment sales, along with other covariates, enables us to uncover the mechanics of the two schemes.

THEORETICAL FRAMEWORK

In this section, we develop a theoretical framework that enables us to gain a better understanding of the subsequent empirical results. We deliberately kept the models simple and created them with our previously described empirical context in mind. Note that the aim here is not to derive an optimal sales compensation scheme. Rather, our goal is to illuminate how the use of bonuses and commissions might involve trade-offs among productivity, multitasking concerns, and distortions induced by timing games.

Incentive Pay and Productivity

To understand the impact of incentive schemes (bonus vs. commissions) on productivity, consider the following stylized model of rational salespeople with differential abilities: For salesperson i exerting effort e_i, sales are given by $y_i = \psi_i e_i$, where $\psi_i > 0$, $e_i \geq 0$, and ψ_i represents the innate "ability" of salesperson i that results in higher output for the same level of effort for a representative with higher ability compared with one with lower ability. We assume it to be distributed continuously with $M \sim [\psi_{min}, \psi_{max}]$, with a higher ψ indicating a more productive salesperson. The firm knows the ability distribution but not the exact ability of each salesperson. We assume throughout that M is strictly increasing with a density m. Our approach is in line with recent work in "personnel economics" (e.g., Lazear 1995, 2000) that has deemphasized the effort–insurance trade-off central to agency models to focus more closely on the role of incentives and salesperson heterogeneity.

The utility of a salesperson with ability ψ_i when he or she earns a wage F and puts in effort e is given by $u(\psi_i;e_i) = F - (\theta e_i^2)/2$, where $\theta > 0$ and represents the cost (of effort) parameter assumed to be common across salespeople.[6] A quota-based bonus scheme is given by the following equation:

$$F_B = \begin{cases} W & \text{if } 0 \leq y_i < Q_B \\ W + B & \text{if } y_i \geq Q_B \end{cases}, \tag{1}$$

where $W > 0$ is the fixed salary irrespective of the output and the salesperson receives a lump-sum bonus $B > 0$ upon reaching or exceeding quota threshold Q_B. An agent will either pick effort 0 (or some positive level of minimum enforced effort) and earn W or exert just enough effort to earn W + B. Because of the heterogeneity in abilities, salespeople would require different levels of effort to reach the bonus quota threshold given by

$$e_{iB} = Q_B/\psi_i. \tag{2}$$

As Equation 2 illustrates, salespeople with greater ability require a smaller level of effort to reach the bonus quota. Therefore, it can be readily verified that that the utility associated with earning a bonus decreases with the decrease in ability. Furthermore, with sufficient heterogeneity in salesperson types, there exists a critical ability threshold above which the salesperson will work more than the minimum required effort and earn the bonus while the rest will simply earn the fixed wage. This threshold level of ability, denoted by ψ_B, can be derived as

$$W + B - \frac{\theta(Q_B / \psi_B)^2}{2} = W.$$

This yields

$$\psi_B^2 = \frac{\theta Q_B^2}{2B}. \tag{3}$$

Assuming that parameters B and Q_B are chosen in such a way that ψ_B lies in the interior of the ability parameter distribution (i.e., at least some people get the bonus), salespeople with abilities $(\psi_B, \psi_{max}]$ respond with more than minimum effort and earn a bonus, whereas salespeople with abilities $[\psi_{min}, \psi_B]$ put in the minimum effort and are paid a fixed wage.

Similarly, a quota-based commission scheme with commission rate $0 < \alpha < 1$ is given by the following equation:

$$F_C = \begin{cases} W & \text{if } 0 \leq y_i < Q_C \\ W + \alpha(y_i - Q_C) & \text{if } y_i \geq Q_C \end{cases}. \tag{4}$$

As before, we can derive the threshold ability of the salesperson who is indifferent between earning a fixed wage and a positive level of commission as

$$\psi_c = \frac{\theta e_c^2 + 2\alpha Q_c}{2\alpha e_c}. \tag{5}$$

A salesperson who (optimally) exerts more than minimum effort solves

$$\underset{e_i}{\arg\max}\ \alpha(y_i - Q_C) - \frac{\theta e_i^2}{2},$$

which yields the optimal effort level

$$e_i^* = \frac{\alpha \psi_i}{\theta}. \tag{6}$$

Equation 6 shows that the equilibrium effort of a salesperson who decides to exert more than the minimum effort level increases in ability. Substituting Equation 6 into Equation 5 yields the following equation:

$$\psi_c^2 = \frac{2\theta Q_c}{\alpha}. \tag{7}$$

As detailed previously, we define a bonus scheme as equivalent to a commission scheme if, for an output Q_B, both schemes result in the same pay. Thus, the condition for the equivalency of two schemes is given by

$$W + B = W + \alpha(Q_B - Q_c) \text{ or } B = \alpha(Q_B - Q_c). \tag{8}$$

[5]We do not suggest that the before–after design is superior to a randomized control–treatment approach. However, unlike a classic randomized control–treatment approach, in our context, it was not possible to isolate the treatment and control groups.

[6]The heterogeneity in salesperson abilities could be equivalently modeled using a heterogeneous cost parameter, with higher ability represented by a lower cost of effort (θ).

Using Equation 8 in conjunction with the other expressions, we obtain our first key result:

Result 1: Under equivalent bonus and commission schemes, the salesperson who is indifferent between earning a positive level of commission and a fixed salary is of lower ability compared with the salesperson who is indifferent between earning a bonus and a fixed salary. In other words, $\psi_c < \psi_B$ (for the proof, see Web Appendix A www.marketingpower.com/jmr_webappendix).

This simple result highlights one source of productivity differences across two equivalent schemes. A positive mass of salespeople who would otherwise engage in minimum effort and earn a fixed salary under the bonus scheme would earn commissions and put in higher effort. In addition, at the higher end of the spectrum, the highest-ability salespeople are no longer constrained by a fixed bonus and tend to invest in higher effort and make more money than they would otherwise make under an equivalent bonus scheme. To demonstrate this formally, we can write the effort as a function of ability under the two schemes as follows:

$$\text{(9a)} \qquad e_{iB} = \begin{cases} 0 & \text{if } \psi_i < \psi_B \\ \dfrac{Q_B}{\psi_i} & \text{if } \psi_i \geq \psi_B \end{cases}, \text{ and}$$

$$\text{(9b)} \qquad e_{iC} = \begin{cases} 0 & \text{if } \psi_i < \psi_C \\ \dfrac{\alpha\psi_i}{\theta} & \text{if } \psi_i \geq \psi_C \end{cases}.$$

Two features are worth noting in Equation 9. First, the participation of agents who respond to incentives has increased under the commissions regime, as evidenced by $\psi_c < \psi_B$ (Result 1). Second, of the agents who respond to incentives under bonuses, the effort is *inversely* related to ability; that is, higher-ability types *need to incur less effort* than lower-ability types to earn a bonus. The opposite holds under commissions: the effort level of agents who earn commissions is *positively* related to ability: higher-ability types *put in more effort* and earn higher commissions than lower-ability types.

Comparing across regimes, the effect on the middle ability types who were already earning a bonus is somewhat ambiguous; for example, see the left-side graph in Figure 2, Panel A, in which there exists a middle segment that puts in less effort than under commissions, whereas in the left-side graph in Figure 2, Panel B, all the ability types put in (weakly) more effort under commissions than under bonuses. However, it is not ambiguous that the increase in effort level of salespeople who were not previously earning bonuses but who now earn commissions is significant.

Although the firm cannot directly observe the salespeople's effort levels, the incentive plan it uses influences the salespeople's efforts, and effort can be inferred from the contract and the distribution of abilities. Thus, the firm can affect the output through the appropriate design of incentive pay. The rational salesperson response to the equivalent bonus and commissions contracts can be expressed as the following output equation:

$$\text{(10a)} \qquad Y_{iB} = \begin{cases} 0 & \text{if } \psi_i < \psi_B \\ Q_B & \text{if } \psi_i \geq \psi_B \end{cases}, \text{ and}$$

$$\text{(10a)} \qquad Y_{iC} = \begin{cases} 0 & \text{if } \psi_i < \psi_C \\ \dfrac{\alpha\psi_i^2}{\theta} & \text{if } \psi_i \geq \psi_C \end{cases}.$$

This output function is plotted in the right-side graphs in Figure 2, Panels A and B. It shows that whereas output remains constant under the bonuses at higher ability levels, it increases dramatically under commissions. Again, the most significant change is for the segment at the lower end, wherein some salespeople who were putting in minimal effort (under bonuses) now respond to incentives and exhibit significant productivity improvement. The middle types' productivity change is somewhat ambiguous, whereas productivity increases for the higher types. In summary, we would expect to observe the largest change in productivity at the lower ability levels.

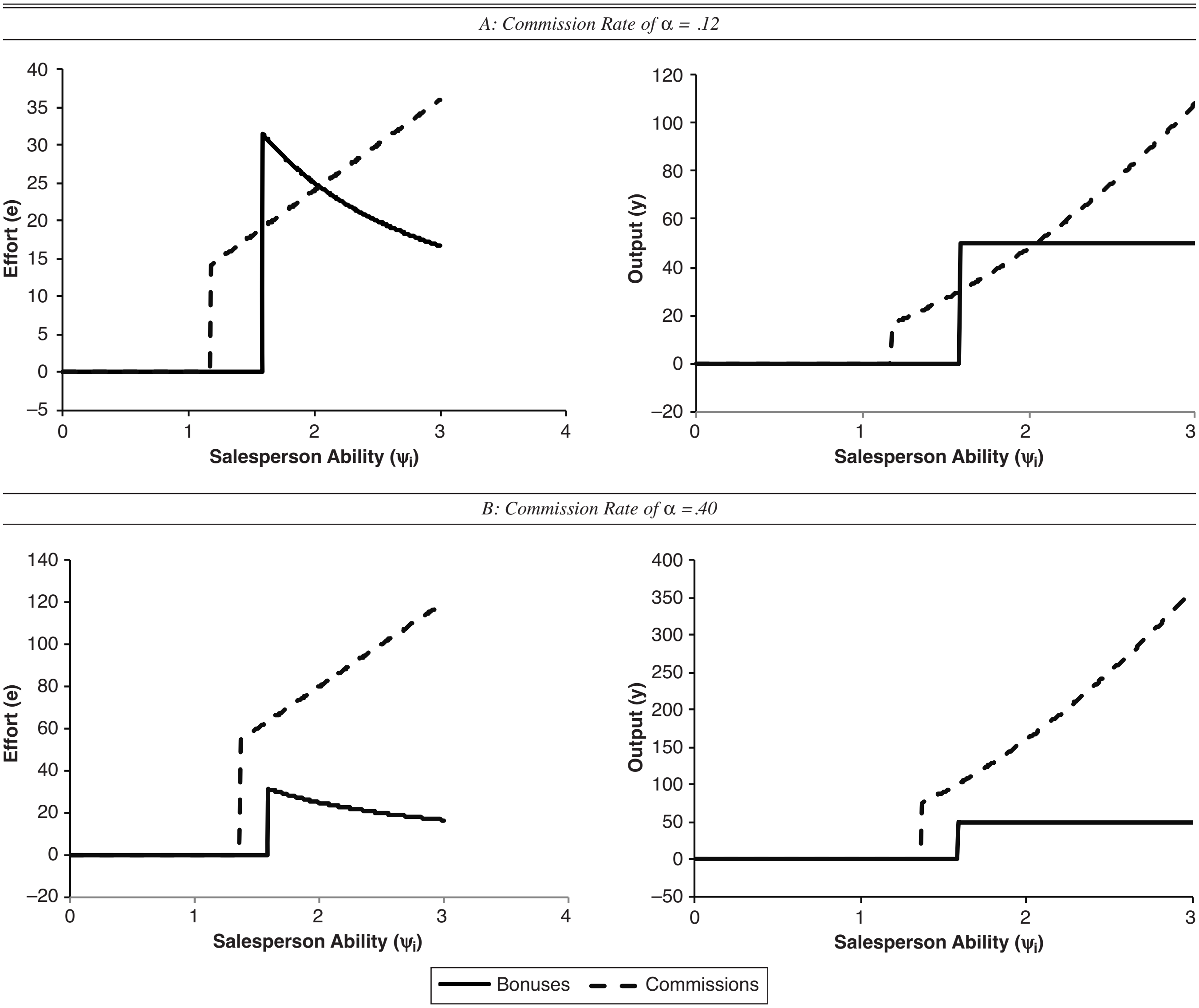
We again emphasize that this result is not meant to convey that commissions are strictly superior to bonuses in terms of productivity. We highlight two sources of possible productivity gains under commissions when two schemes are equivalent (even if suboptimal): low-ability agents putting in more than minimum effort for a chance to earn nonzero commission and high-ability agents putting in higher effort because there is no earning ceiling.

Multitasking and Timing Games

The model previously presented includes one task and a deterministic relationship between effort and output. However, as we described previously, this firm's salespeople mainly undertake two tasks that are not equally influential in determining immediate sales (which are the basis for determining incentive pay in both the pay plans). Furthermore, they could pull in and/or push out sales. We aim to understand these issues with a combination of a standard multitask model (Holmstrom and Milgrom 1991) and a simple timing game model (Oyer 1995).

Consider a salesperson engaged in two activities over two quarters, each consisting of two periods (We acknowledge that real-world quarters consist of three months, but we assume two months per quarter to simplify the exposition of our model while still capturing the sales variation present within the months of a quarter.) To match this model to our empirical setting, we assume that the firm requires a salesperson to engage in two activities, denoted as "doctor visits" and "pharmacy visits," respectively. Furthermore, in the second period of the first quarter, the salesperson can possibly "game" the sales below or above the "natural" sales level, which has consequences for sales in the third period (the first month of the second quarter). The salesperson is paid on total observable output in a quarter. In any given period t (month), observed output (sans gaming) is given by the equation $y_t = h_d T_{dt} + h_p T_{pt} + \phi$, where T_{dt} and T_{pt} denote the time spent on doctor visits and pharmacy visits (respectively) in period t and h_d and h_p are the positive marginal products of these activities on observed output. The time spent proxies effort substitution across these activities, and ϕ represents random noise distributed symmetrically around mean zero. The value added to the firm arising from the time spent on these activities is given by the equation $\nu_t = f_d T_{dt} + f_p T_{pt} + \varepsilon$, where f_d and f_p are the positive marginal

Figure 2
EFFORT AND OUTPUT UNDER EQUIVALENT BONUS AND COMMISSION PLANS



Notes: The parameter values we used in constructing these figures are B = 5, Q_B = 50, θ = .01, and ψ ~ u [0, 3]. The parameter that is different across Panels A and B is the commission rate of α = .12 and .40, respectively.

products of these activities on firm value and ε represents mean zero random noise with properties similar to ϕ as described previously.

Furthermore, $f_d > f_p$ and $h_d > h_p$, setting up a classic multitasking problem wherein the activities that enhance the firm's welfare are valued less by the salesperson because of their lesser impact on observable output. The salesperson cannot be compensated directly on ν_t because it is virtually impossible to quantify an individual agent's contribution to the firm's value. The total time spent on the two activities, T, is fixed in a given period t (salespeople are assumed to work a certain fixed number of hours in a month), and $T = T_{dt} + T_{pt}$ with the minimum number of doctor visits in any period fixed at $T_{d\ min}$. In the bonus plan, in quarter 1, the salesperson is compensated W + B if $y_1 + y_2 \geq Q_B$; otherwise, he or she gets W. Similarly, in quarter 2, if $y_3 + y_4 \geq Q_B$, he or she makes W + B; otherwise, he or she gets W.[7] In the commissions plan, the salesperson makes $W + \alpha (y_1 + y_2 - Q_c)$ if $y_1 + y_2 \geq Q_c$ with $0 < \alpha < 1$; otherwise, he or she earns W. The same scheme applies in quarter 2. Equivalence requires $\alpha (Q_B - Q_C) = B$ and, trivially, $Q_B \geq Q_C$.

A well-established literature stream has shown that salespeople often game output in response to high-powered incentives (Jensen 2003; Misra and Nair 2011; Steenburgh 2008). Drawing from Oyer (1995), we model this finding by allowing that in period 2 of quarter 1, the salesperson

[7]We are abstracting away from so-called ratcheting effects by assuming a constant quota over time (Gibbons 1987). Admittedly, this is not an innocuous assumption, but we have both institutional and empirical evidence to suggest that ratcheting is unlikely to be an important strategic consideration among salespeople in our setting. For details, see the "Discussion" section and Web Appendix B (www.marketingpower.com/jmr_webappendix).

could potentially game the system by either pulling in sales from period 3 or pushing out sales to period 3. Both these activities are costly (to the firm). We model this gaming through a variable $0 < \lambda < 2$, where $\lambda = 1$ implies a "natural" level of sales, $\lambda > 1$ indicates pull-in, and $\lambda < 1$ indicates push-out. Finally, we assume that when a salesperson is indifferent to carrying out either doctor or pharmacy visits, he or she focuses on doctor visits. Figure 3 gives a summary of the time line of this game. The critical variable that influences a salesperson's gaming strategy and activity substitution is the realized sales y_1 that translates into the state variable S_{Q1B} (distance to quota).

Given this framework, the salesperson's observed output and contribution to firm value in each of the four periods can be written mathematically. We have relegated these derivations to Web Appendix A (www.marketingpower.com/jmr_webappendix) and give the key mathematical expressions in Table 1. It is worth emphasizing that in this formulation, we aim to understand salesperson behavior under a given bonus scheme and an equivalent commission scheme while abstracting away from the firm's problem.

We use the salesperson's problem laid out in Table 1 to gain insights into the choices made in the last month of a quarter and the impact of these choices in the first month of the subsequent quarter under bonus and commission regimes, respectively. We collect two key empirically testable results, which we discuss subsequently. Web Appendix A (www.marketingpower.com/jmr_webappendix) details an example of the algebra involved in deriving these results. These two results show that when the quota is too far (and therefore unlikely to be reached), salespeople exhibit similar behavior across bonus and commission regimes; in contrast, when the quota has been met, salesperson behavior is dramatically different across the two regimes in terms of timing games and multitasking distortions.

Result 2: If, at the end of period 1, the quota for bonus is too far (we label this condition "FAR" in our empirical analysis), the salesperson will push sales out to period 3, and the focus will shift (weakly) to doctor visits, resulting in attenuation of multitasking concerns. The same results occur if the quota for commissions is too far at the end of period 1.

Result 3: If, at the end of period 1, the quota for bonus has been achieved (we label this condition "EXCEEDED" in the empirical analysis), the salesperson will push sales out to period 3 and will focus largely on doctor visits, resulting in the attenuation of multitasking concerns. However, if the quota for commissions has been achieved, the salesperson will pull sales in from period 3 and will focus largely on pharmacy visits while keeping doctor visits at the minimum level, resulting in the amplification of multitasking concerns.[8]

The intuition behind Result 2 is simple: there is little chance that the salesperson will reach the quota in either regime, so the marginal return to pull-in is almost zero. Therefore, the salesperson will push sales out to improve his or her chances of reaching/exceeding the quota in the next quarter. Because doctor visits dampen the current output but returns to the current output decrease are nonnegative, we expect salespeople to focus weakly on doctor visits. In Result 3, we observe the clearest distinction between the effects of the two

[8]The terms "near," "too far," and "far" (stretch) are defined more precisely on pp. 7–8 of Web Appendix A (www.marketingpower.com/jmr_webappendix) with reference to the model. The corresponding empirical measures are detailed in the "Empirical Analysis" section.

Figure 3

SEQUENCE OF EVENTS IN MULTITASKING AND TIMING GAME MODEL

Table 1

SALESPERSON OPTIMIZATION PROBLEM

	Probability of Reaching Quota in Quarter 1	*Expected Incentive Quarter 1*	*(Expected) Probability of Reaching Quota in Quarter 2*	*Expected Incentive Quarter 2*	*Cost of Gaming*
Bonus	$1 - G[(S_{Q1B}/\lambda) - h_dT - (\Delta h)T_{p2}]$	B	$1 - G[S_{Q2B} - h_dT - (\Delta h)T_{p4}]$	B	$\mu(\lambda - 1)^2 D$
Commissions	$1 - G[(S_{Q1C}/\lambda) - h_dT - (\Delta h)T_{p2}]$	$\alpha\{\lambda[h_dT + (\Delta h)T_{p2}] - S_{Q1C}\}$	$1 - G [S_{Q2C} - h_dT - (\Delta h)T_{p4}]$	$\alpha[h_dT + (\Delta h)T_{p4} - S_{Q2C}]$	$\mu(\lambda - 1)^2 D$

Notes: The salesperson optimization problem can be represented by arg $\max_{(\lambda, T_{p2})}$[(Probability of reaching quota in Quarter 1) × (Expected incentive Quarter 1)] + β[E(Probability of reaching quota in Quarter 2) × (Expected Incentive Quarter 2)] – [Cost of gaming].

schemes. Under bonuses, the salesperson has no incentive to improve output after reaching the quota; indeed, he or she has every incentive to push sales out to period 3 to maximize the probability of reaching quota in the next period as well. Yet the salesperson will readily focus on doctor visits even though pharmacy visits are more impactful in the realization of observable output, because the marginal return to increased output is zero after reaching the quota. Under the commission scheme, the marginal return to increased output is positive, and the salesperson will therefore largely focus on pharmacy visits to maximize current sales. The salesperson will also pull sales in from period 3 because he or she may not reach the quota in the next period; having reached the quota in this period, the salesperson wants to sell as much as possible. To recap, a realistically achievable bonus plan tends to amplify distortions induced by timing games, whereas an equivalent commission plan exacerbates multitasking concerns. Table 2 summarizes the key results from our theoretical framework.

EMPIRICAL ANALYSIS

Data Description

Our data include an almost complete history of incentive plans and payments, monthly output, and quarterly quotas for 458 territories for a period of three years, from April 2007 (the beginning of the fiscal year) to March 2010. The data also include average daily doctor and pharmacy visits at the monthly level. The intervention in October 2008 marks the beginning of the third quarter of the 2008–2009 fiscal year. The preintervention period of 18 months (under the bonus plan) begins in April 2007 and ends September 2008, and the postintervention period of 18 months for these same territories (under the commission plan) begins in October 2008 and ends March 2010. In all, the data include more than 14,000 monthly observations on both output and visits.

The descriptive statistics reported in Tables 3 and 4 show large differences in quotas across territories; in other words,

Table 2
RESULTS FROM ANALYTICAL MODEL

Result	*Productivity*	*Heterogeneity in Productivity Increase*	
1	Bonus < commissions	↑ Low ability > ↑ high ability Medium ability ↔	
	Timing Games	*Too Far from Quota (FAR)*	*Exceeds Quota (EXCEEDED)*
2a–3a	Bonus	Push sales out	Push sales out
2c–3c	Commissions	Push sales out	Pull sales in
	Multitasking	*Too Far from Quota (FAR)*	*Exceeds Quota (EXCEEDED)*
2b–3b	Bonus	Doctor ↑ Pharmacy ↓ (small effect)	Doctor ↑ Pharmacy ↓
2d–3d	Commissions	Doctor ↑ Pharmacy ↓ (small effect)	Doctor ↓ Pharmacy ↑

Table 3
DATA DESCRIPTION

Variable	*Definition*	*M*	*SD*
NewPlan	A dummy variable equal to 1 if sales group is on the commission plan during that quarter/month	.51	
GroupSize	Total members in the sales group	2.17	3.05
Quarterly productivity	Quarterly revenues generated per salesperson	27.60	15.63
Monthly productivity	Monthly revenues generated per salesperson	9.26	5.63
Incentives	Quarterly incentives earned by a salesperson	.24	.42
Net revenues	Quarterly revenues minus total incentives paid out to sales group	57.15	72.56
Industry sales	Total quarterly industry sales	210.66	40.96
Target	Quarterly sales quota	62.37	76.49
Sales History Variables			
EXCEEDED	If quarterly performance to date is ≥1, this variable takes a value of 1 during the last month of the quarter and 0 otherwise	.02	
NEAR	If quarterly performance to date ≥2/3 but <1, this variable takes a value of 1 during the last month of the quarter and 0 otherwise	.10	
STRETCH	If quarterly performance to date ≥1/3 but <2/3, this variable takes a value of 1 during the last month of the quarter and 0 otherwise	.17	
FAR	If quarterly performance to date ≥0 but <1/3, this variable takes a value of 1 during the last month of the quarter and 0 otherwise	.02	
POST EXCEEDED	Takes value of 1 if EXCEEDED = 1, and 0 if otherwise	.02	
POST NEAR	Takes value of 1 if NEAR = 1, and 0 if otherwise	.08	
POST STRETCH	Takes value of 1 if STRETCH = 1, and 0 if otherwise	.14	
POST FAR	Takes value of 1 if FAR = 1, and 0 if otherwise	.02	

Notes: We disguised productivity, net revenues, targets, and incentives by multiplying actual values by a single common positive number. Furthermore, the currency is suppressed for confidentiality.

Table 4
KEY VARIABLES BY INCENTIVE PLAN

	Bonus Plan		Commissions Plan	
Variable	*M*	*SD*	*M*	*SD*
Monthly productivity	8.99	5.63	9.51	5.61
Quarterly productivity	26.23	15.97	28.87	15.20
Quarterly net revenues	51.87	64.42	62.07	79.08
Quarterly incentives	.14	.39	.34	.43
GroupSize	2.06	2.75	2.23	3.21
Quarterly target	70.44	82.13	57.95	75.12
10th percentile quarterly target	18.47		13.14	
90th percentile quarterly target	132.94		114.89	

there are significant differences in sales potential across territories. Figure 4 plots total monthly sales. Visually, revenues appear to have increased postintervention. Substantial seasonal variation is also evident. Table 3 presents summary statistics for revenues, productivity (defined as revenues per salesperson per quarter), and incentive pay earned by the compensation plan. Consistent with Figure 4, the numbers in Table 4 indicate that revenues and productivity are higher in the intervention period. Figure 5 presents a plot of industry sales over time during the observation period, while Figure 6 shows total monthly sales of the firm including the 5th and 95th percentile territories.

Linking Theory to Empirical Application

We elaborate on the suitability of our empirical context and design to the task of assessing the validity of our theory results here. Consider the results in turn.

Result 1 on productivity relies on two key elements of the model: (1) heterogeneity in the innate ability of the salespeople and (2) equivalence of the two schemes. Because we have already described the operationalization of the equivalent scheme, we focus on the first element here. Heterogeneity, combined with chronic/persistent ability, yields the outcome wherein the lower types respond to incentives under commissions ($\psi_c < \psi_B$). Does our context yield data with sufficient heterogeneity in ability but also persistence in this ability over time, specifically across bonus and equivalent commissions? Given that we observe the productivity of each sales group across multiple quarters under bonuses and equivalent commissions, we can readily test this identification condition. Using the productivity Equation 11, we estimate the fixed effects associated with each sales group separately for bonuses and commissions (by running different regressions for each regime). Figure 7 plots these fixed effects; there is heterogeneity in ability estimates as well as ability persistence across the regimes (the correlation coefficient between the fixed effects is .742).

Furthermore, while Result 1's prediction of salespeople responding more to incentives under commissions is supported by the quota attainment numbers (under the bonus regime, only 14% of salespeople met their quarterly quota, whereas 60% did so for the commission regime), our design also generates additional evidence for the mechanism. Although we do not have direct measures of effort and ability, the efforts can be linked to output through Equation 10, and as we previously demonstrated, we can infer the salesperson ability through repeated observations. Thus, the direction of the productivity change at different ability levels across bonuses and commissions in our experimental setup enables us to assess the change in effort, which we accomplish through quantile regressions, as explained in the "Heterogeneous Intervention Effect" subsection.

Figure 4
TOTAL MONTHLY SALES OVER TIME

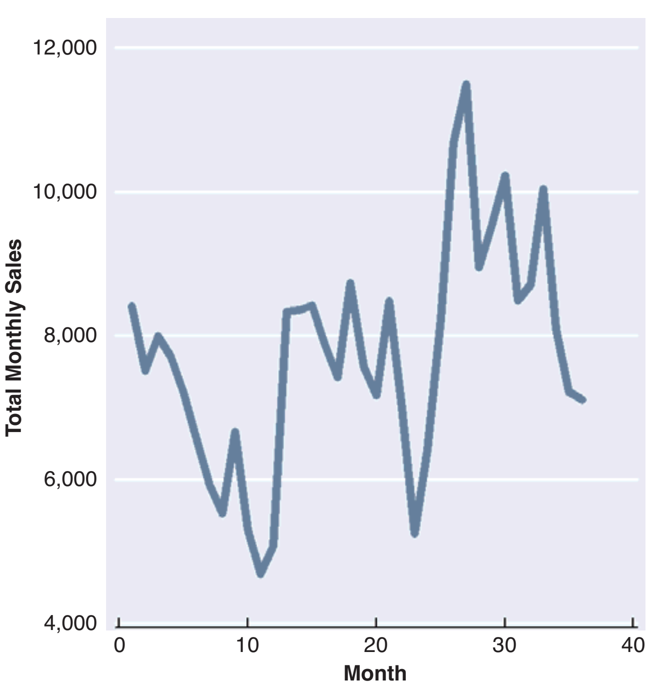


Figure 5
MONTHLY INDUSTRY SALES OVER TIME

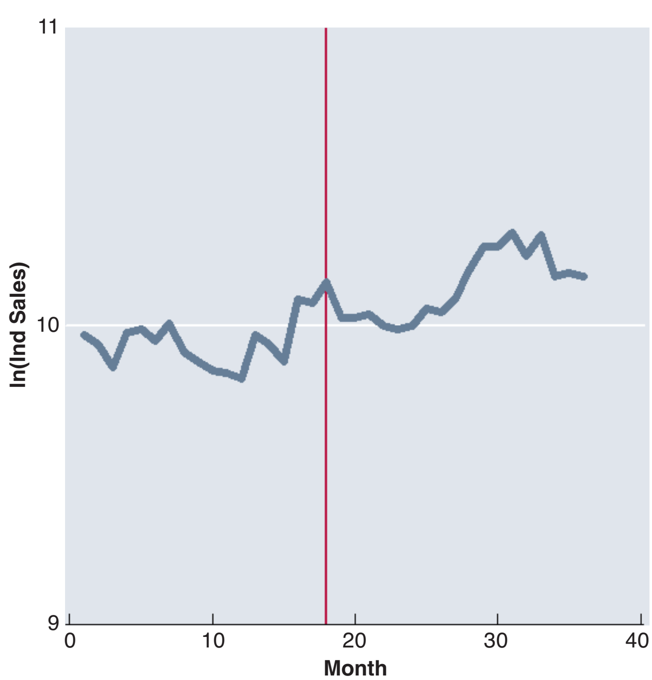
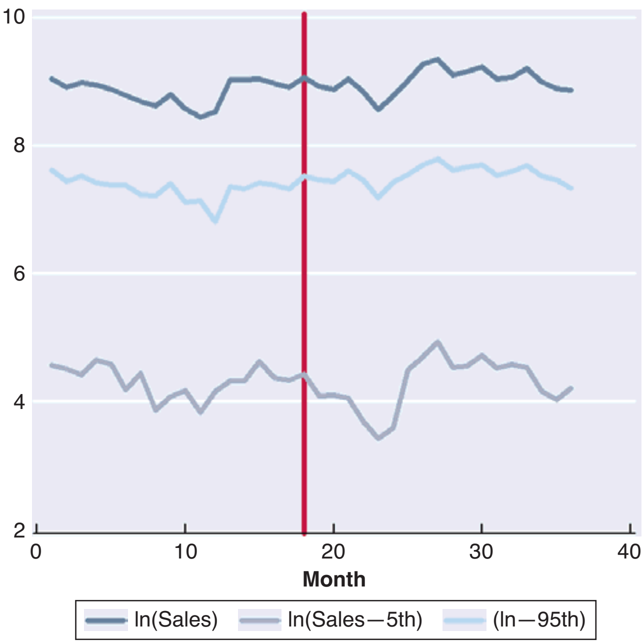


Figure 6
TOTAL MONTHLY SALES OVER TIME



Our multitasking results hinge on the institutional observation that pharmacy (doctor) visits do (do not) translate into immediate sales gain. We arrived at this crucial working assumption through extensive conversations with the management as well as observation during field visits. We are also able to test the empirical validity of this crucial assumption by running a regression of monthly sales against pharmacy and doctor visits and other controls (see Table B4 in Web Appendix B at www.marketingpower.com/jmr_webappendix). Indeed, we find that the coefficient of pharmacy visits is positive (.0603, $p < .01$), whereas the coefficient of doctor visits is negative (–.007, $p = .659$) but statistically not significant.[9] This is in accordance with our two-task multitasking model outlined previously and discloses the appropriateness of our empirical setting to the theoretical insights laid out in Results 2 and 3.

Intervention Effect on Productivity

Figure 8 plots the kernel density of quarterly productivity (defined as unit sales per salesperson). From this figure, it is evident that the commission distribution lies to the right of the bonus distribution. Furthermore, the peak value of the density function in the bonus plan is lower than that of the commission plan. There is also more concentration of productivity around the modal value under commissions than under bonuses. Table 4 describes mean pre- and postintervention productivity levels as 26.23 and 28.87, respectively, which provides directional evidence that the intervention increased productivity. We estimate the following model to analyze the intervention effect more formally:

$$(11)\quad \log(y_{it}) = \alpha_0 + \sum_i \alpha_i S_i + \sum_t \gamma_t YM_t + \beta NewPlan_t + \delta_1 \log(QtrTarget_{it}) + \delta_2 GroupSize_{it} + \varepsilon_{it},$$

where

y_{it} = revenue per salesperson generated in territory i in period t (e.g., July 2008);
S_i = dummy for territory i that controls for the unobserved ability/territory effect of the sales group;
YM_t = year–month dummy for period t (e.g., Jan 2008), which allows month dummies to be estimated from the within-month variation across the 458 territories and provides a stringent control for unobserved differences in demand/competition/business environment between the before and after periods[10];
$NewPlan_t$ = intervention dummy set to 0 for pre-intervention periods;
$QtrTarget_{it}$ = quarterly quota for territory i for period t that provides observed control for the differences across territory potential over time; and
$GroupSize_{it}$ = number of salespeople in territory i in period t, which provides another observed control for the territory effects over time. The disturbance term ε_{it} is clustered at the territory level.

Identification of productivity estimate. The main identification assumption for the productivity estimate by plan is the exogeneity of the intervention (the compensation switch was designed as a quasi-experiment). Because we do not have a control group, the use of month–year and territory dummies along with the measures of quarterly quota and sales group size is useful to control for unobserved and observed territory, time, and other market-level heterogeneity. A few issues are worth noting. Quarterly quota could potentially be endogenous, and although the covariates and unobservable heterogeneity control alleviate this concern somewhat, we also estimate our productivity specification without including the quota variable (see Table 5). This also alleviates the concern that the quota systematically varies with scheme change and that its inclusion might create biased estimates of the new plan.

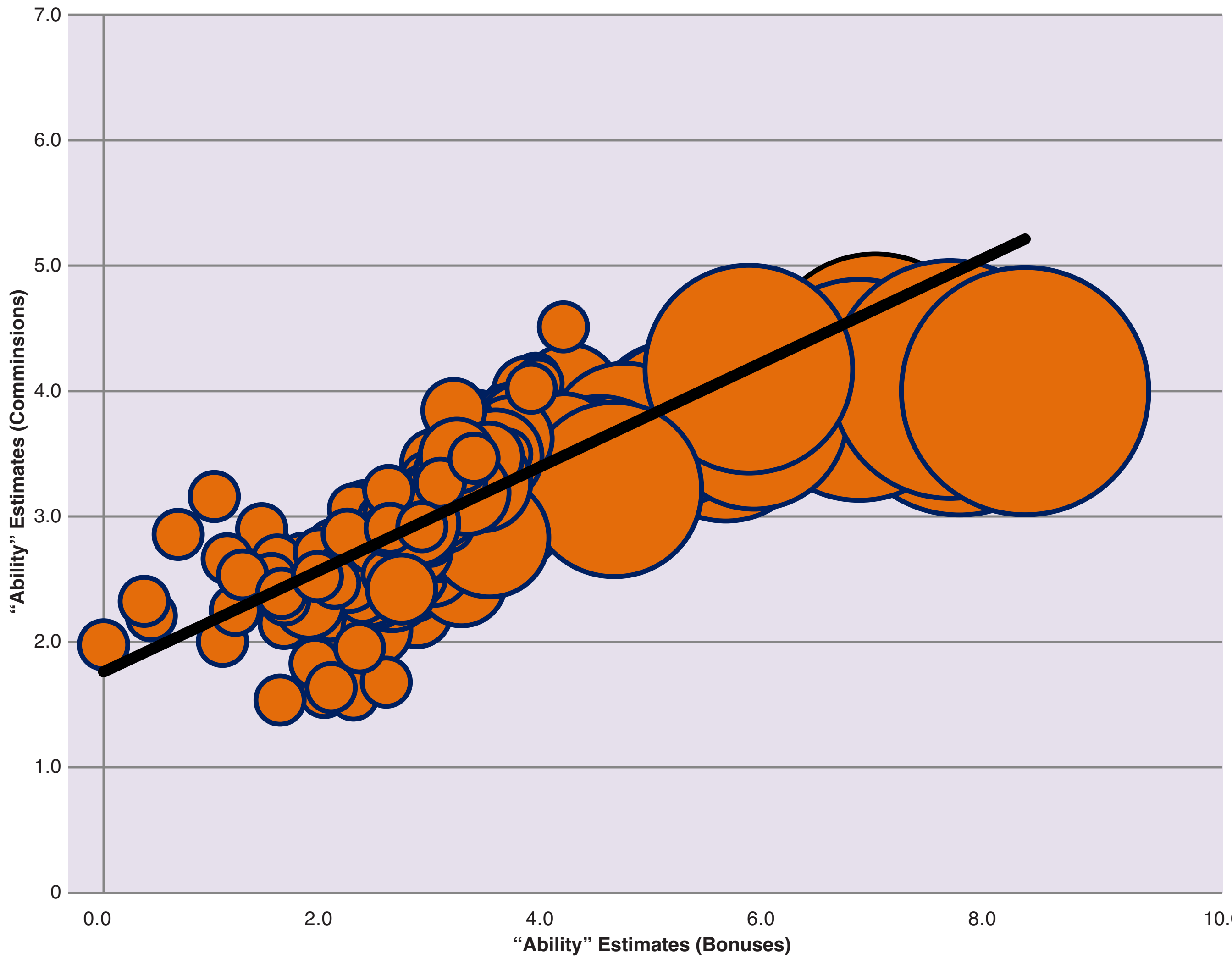
Selection, terminations, and group-size variation. We have interpreted the territory dummy as being a control for a sales group's underlying ability and territory potential. This interpretation is problematic if the team composition changes because of terminations and hires (for further details, see Web Appendix B at www.marketingpower.com/jmr_webappendix). We accounted for turnover and reran the entire analysis by constructing fixed effects at the territory–team level rather than just at the territory level. For example, if a sales group had 12 members during the first 14 months, 10 members during the next 6 months (due to people leaving), and 13 members during the last 16 months (due to 3 new members joining), we gave each of these three territory–team pairs three distinct identifications and estimated them as distinct fixed effects in the analysis. In Table 5, we present the results of all the specifications of

[9]If the total number of visits is fixed (as in the theory model), some may question how the effect of doctor and pharmacy visits are separately identified. We are able to make the identification because we do observe some variation in total visits across groups as well as over time.

[10]We are grateful to an anonymous referee for suggesting this control. Our results are robust to noninclusion of this set of controls.

Figure 7

ABILITY ESTIMATES ACROSS BONUSES AND COMMISSIONS



Notes: The size of each circle represents the average size of the sales group during the period of observation, and the center of the circle is the associated fixed effect estimate. The black solid straight line represents the linear trend line. We estimated each territory fixed effects pair by running Equation 11 for each regime. We have normalized the smallest fixed effect to zero.

Equation 11 using these new territory–team controls as well. The results are fairly similar across the specification with the territory fixed effects (Columns 1, 2, and 3) and team–territory fixed effects (Columns 4, 5, and 6). We also reran the analyses by dropping all the territories that experienced any turnover and produced similar results, which we present in Web Appendix B (www.marketingpower.com/jmr_webappendix) along with other robustness results.

Column 3 of Table 5 shows the estimate of our intervention effect on productivity. Evaluated at the mean, an average territory is 23.9% more productive after the intervention. When accounting for turnover, the new plan productivity estimate shows a similar productivity improvement of 24.9% (see Column 6 of Table 5). Because these estimates are the net estimated effects after controlling for calendar month–year and unobserved territory differences, it is reasonable to attribute them to increased salesperson effort. (The coefficients of all but one period dummy are significant, indicating the presence of seasonality as well as a time trend.) The noninclusion of the quarterly target results in a higher value of this estimate (see Columns 2 and 5 of Table 5), whereas noninclusion of the month–year dummies results in a lower value of this estimate (see Columns 1 and 4 of Table 5). Overall, according to Table 5 with different specifications, the most conservative estimate is a productivity improvement of approximately 22% as a result of the switch.

Heterogeneous Intervention Effect

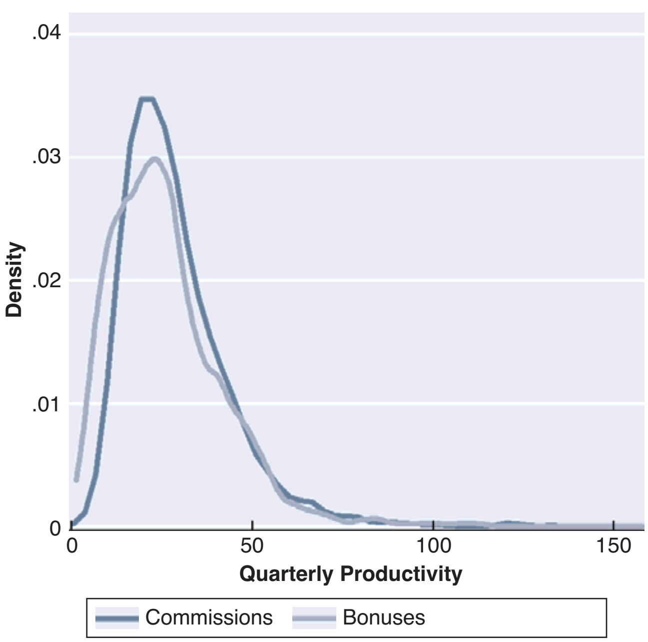
Recall that we predicted differential responses to the plan change across agents of different abilities from our model. We use the following quantile regression specification to estimate these differences:

Table 5
MONTHLY PRODUCTIVITY RESULTS

	Turnover Not Accounted			Turnover Accounted		
	1	*2*	*3*	*4*	*5*	*6*
NewPlan	.222** (.008)	.558** (.0287)	.239** (.0300)	.238** (.008)	.576** (.0287)	.249** (.0301)
GroupSize	–.091** (.009)	–.0804** (.009)	–.0863** (.009)	–.048** (.008)	–.0120 (.008)	–.0136* (.008)
log(QtrTarget)	.420** (.012)	—	.178** (.0172)	.432** (.012)	—	.192** (.0175)
Constant	.534** (.0492)	1.73** (.028)	1.39** (.061)	.391** (.047)	1.52** (.028)	1.19** (.061)
Group effects	Included	Included	Included	Included	Included	Included
Month–year effects	Not included	Included	Included	Not included	Included	Included
Observations	14,499	14,499	14,499	14,499	14,499	14,499
R-square	.597	.656	.659	.603	.662	.665
Sales groups (clusters)	458	458	458	528	528	528

*$p < .1$.
**$p < .01$.
Notes: Standard errors in parentheses. Dependent variable = log (monthly productivity). Columns 1, 2, and 3 use territory fixed effects as the group effects, and Columns 4, 5, and 6 report the results in which team composition (turnover) is accounted for and each new team in a territory is assigned a unique dummy to create the "team–territory fixed effects."

Figure 8
QUARTERLY PRODUCTIVITY



$$\text{(12)}\quad \text{Quant}_\theta[\log(y_{it})\,|\,.] = \alpha_{0\theta} + \sum_t \gamma_{t\theta}\text{YM}_t + \beta_\theta\text{NewPlan}_t + \delta_{1\theta}\log(\text{QtrTarget}_{it}) + \delta_{2\theta}\text{GroupSize}_i + \varepsilon_{it},$$

where all variables are as previously defined. Columns A–E in Table 6 report the simultaneous estimates of Equation 12 at different quantiles indexed by θ. The results illustrate that the average intervention effect reported here masks substantially different changes throughout the conditional distribution of log productivity.

Our intervention increases the productivity of the 10th quantile territory by 41.5%, the 25th quantile by 37.0%, the 75th quantile by 15.1%, and the 90th quantile by 20.0%. Overall, the quantile regression estimates confirm that the switch to the commission plan has differential effects on the productivity of the best and worst sales groups, with the greatest impact on the least productive groups. In addition, the variance in output decreases postintervention because lower-productivity groups experience a significant increase in productivity and move closer to the productivity of higher-performing groups.

Timing Games

Table 2 summarizes the predicted effects on timing games. Parenthetically, we clarify that our timing games effects do not imply that a salesperson at a certain distance from the quota will not increase effort to reach the quota. We argue only that a part of any observed acceleration in sales as he or she approaches the quota could be explained by the salesperson pulling in sales from the future. As such, the more rigorous test of our prediction comes from the distinctive predictions based on how far a salesperson is from the quota in the specific regime in question. Evidence of pull-in (push-out) from the future in the final month of a quarter is not simply an observed increase (decrease) in sales but whether, in addition, there is a statistically significant decrease (increase) in sales in the first month of the following quarter. Following Steenburgh (2008), we define categorical variables that capture the impact of past performance on a salesperson's current output. We constructed variables using performance to date against quota immediately before the final month of the quarter. We define performance to date as the ratio of the cumulative revenue produced in the first two months of the quarter to the quota that must be met in that quarter. Then, we used this performance-to-date variable to create the categorical variables EXCEEDED, NEAR, STRETCH, and FAR for the last month of the quarter. We also created the variables POST EXCEEDED, POST NEAR, POST STRETCH, and POST FAR in the month following the quarter-ending month. Table 3 provides detailed definitions.

As we mentioned previously, if salespeople only vary their efforts over months within quarters, spikes or dips in revenue production could occur in the last month of the quarter but not in the month immediately following the quarter. However, if salespeople are playing timing games,

Table 6
EFFECT OF COMMISSION PLAN ON PRODUCTIVITY (QUANTILE SUR REGRESSIONS)

Variables	*A* *10th*	*B* *25th*	*C* *50th*	*D* *75th*	*E* *90th*
NewPlan	.415** (.098)	.370** (.063)	.219** (.047)	.151** (.040)	.200** (.053)
GroupSize	–.169** (.007)	–.147** (.005)	–.115** (.004)	–.092** (.002)	–.094** (.001)
log(QtrTarget)	.857** (.016)	.696** (.012)	.576** (.011)	.524** (.010)	.567** (.012)
Constant	–1.85** (.095)	–.892** (.067)	–.117* (.049)	.426** (.035)	.529** (.064)
Month–year effects	Included	Included	Included	Included	Included
Observations	14,499	14,499	14,499	14,499	14,499
(Pseudo) R-square	.289	.2607	.2459	.2454	.2550

*$p < .05$.
**$p < .01$.
Notes: Standard errors (bootstrapped) in parentheses.

spikes and dips in revenue in the last month of a quarter will be followed by dips and spikes, respectively, in the first month of the next quarter. We can therefore use the coefficients on the POST variables to infer whether timing games are present. We estimate the following model of revenue in territory group i in period t with the following equation:

$$(13)\quad \log(r_{it}) = \alpha_0 + \sum_i \alpha_i S_i + \sum_m \nu^m \Gamma_{it}^m + \sum_{post-m} \nu^{post-m} \Gamma_{it}^{post-m} + \delta_1 \log(QtrTarget_t) + \delta_2 GroupSize_i + \varepsilon_{it},$$

where r_{it} is the average revenue produced in territory i in period t and S_i, $QtrTarget_t$, and $Groupsize_i$ are all as defined previously; $\Gamma_{it}{}^m$ consists of the EXCEEDED, NEAR, STRETCH and FAR dummy variables describing sales to date in the last month of a quarter in territory i in period t; and Γ_{it}^{Post-m} consists of the POST EXCEEDED, POST NEAR, POST STRETCH, and POST FAR dummy variables describing the state in the first month of the subsequent quarter.

We estimated the preceding regression separately on data from the bonus and commission plans. Column 1 in Table 7 shows evidence of timing games under the bonus regime. Specifically, both the EXCEEDED and the FAR coefficients are negative and significant (–.417, $p < .01$, and –.316, $p < .01$, respectively). This finding suggests that sales groups might be pushing sales out if (1) they have exceeded quota or (2) they expect that they cannot make the quota. In reviewing the POST variables, we observe that both coefficients on the POST EXCEEDED and POST FAR variables are positive and significant (.308, $p < .01$, and .163, $p < .01$). Overall, this finding provides confirmation for Results 3a and 2a, respectively, in Table 2 that sales groups play timing games under the bonus plan; both sales groups that have exceeded quota and sales groups that are far from achieving quota reduce effort in the final month of the bonus period and push out sales to the next period. In addition, note that the POST EXCEEDED and POST FAR estimates are lesser in absolute value than the EXCEEDED and FAR estimates, suggesting that in the final month, they use effort (reduction) and push-outs in conjunction. The estimates for NEAR and STRETCH coefficients are positive and significant, suggesting that sales groups that are close to quota might be pulling sales in (and increasing effort) to achieve quota and earn incentives. The POST STRETCH estimate is negative and significant, suggesting that sales groups are pulling in sales from the next period to meet quota and earn incentives in line with Result 3a. In contrast, the POST NEAR estimate is not statistically significant.

Table 7
SALES VARIATION ACROSS TIME

Variables	*A* *Bonus Plan*	*B* *Commissions Plan*
EXCEEDED	–.417** (.085)	.279** (.027)
NEAR	.147** (.023)	.096** (.013)
STRETCH	.051** (.012)	.053** (.015)
FAR	–.316** (.041)	–.192* (.100)
POST EXCEEDED	.308** (.070)	.046* (.025)
POST NEAR	.026 (.038)	–.001 (.012)
POST STRETCH	–.121** (.014)	–.005 (.014)
POST FAR	.163** (.041)	.073 (.078)
GroupSize	.0256 (.031)	.018** (.005)
log(QtrTarget)	.379** (.018)	.508** (.017)
Constant	.899** (.089)	.647** (.062)
Fixed effects	Included	Included
Observations	7,107	7,392
Clusters	444	438
R-square (overall)	.8061	.8530

*$p < .1$.
**$p < .01$.
Notes: Standard errors in parentheses, clustered by territory.

Steenburgh (2008) reports an absence of timing games in his analysis of the durable goods office equipment sector and suggests that the finding is likely a function of his particular institutional context. Our context differs greatly from Steenburgh's. For example, in his context, order-booking involves long negotiation periods, management involvement from both sides, bank credit lines, and working capital positions, among other things. Coupled with the sheer dollar amount of the orders, these factors make it difficult to move the order booking date forward or backward. In contrast, our selling context is much more fragmented, with a much smaller monetary value for each transaction, rendering it much more amenable to manipulation.

Column 2 in Table 7 reports results under the commission regime. In the FAR condition, we expect that even under the commission regime, the salesperson would push sales out to the next quarter. The estimates of the FAR and POST FAR variables are –.192 and .073, respectively, both directionally consistent with our Result 2c in Table 2; however, the latter variable is not statistically significant. The estimate of

EXCEEDED is positive and statistically significant (.279, $p < .01$), but the analogous POST estimate is only marginally significant. Therefore, although we predict a weak pull-in, the evidence seems to point toward a limited amount of timing games. In other words, under the commission regime, we find mostly directionally consistent but statistically weak evidence for Results 2c–3c in Table 2. The near absence of timing distortions in the commission regime is intriguing and calls for deeper theoretical and empirical investigation.

Multitasking Distortions

Recall that our salespeople were encouraged to visit physicians and pharmacies. These doctor visits were considered to have little short-term impact on sales but to have long-term consequences for the firm. The firm requires each territory to make at least 80% of their doctor visits in a quarter. If sales groups fall below this level, the firm imposes a penalty, which is usually a fraction of the quarterly incentives earned, if any. In contrast, the firm does not place any corresponding requirements for pharmacy visits. As we explained previously, because pharmacy visits are closely linked to losses in short-term sales from insufficient stocking, salespeople will undertake necessary effort to ensure that pharmacies stock and push products even without any defined minimum call levels.

Our call report data show that most salespeople make between 14 and 16 visits per day (doctor and pharmacy visits), with the 80% enforcement limit falling close to 7.5 doctor visits per day. Table 8, Panel A, reports average daily doctor and pharmacy visits by regime. Under commissions, average daily doctor visits decrease from 10.08 to 7.93, but average daily pharmacy visits increase from 4.92 to 6.74. These averages are statistically different at the 5% level. Apparently, attention shifted to tasks that are directly related to compensation and away from tasks that are nevertheless important to the firm. Yet these raw averages hide a much more nuanced story, which Table 8, Panel B, illustrates. As the "Doctor Visits" numbers show, it is evident that the salespeople who managed to reach quota were just fulfilling the minimum level of doctor visits under both the bonus and commission regimes (8.02 and 8.50, respectively) before reaching the quota. This is in line with Corollary 1 (Web Appendix A; www.marketingpower.com/jmr_webappendix). However, after the quota has been achieved, the salesperson under the bonus regime shifts his or her attention largely to doctor visits ($12.98 > 8.02$, $p < .01$), whereas attention shifts away from doctor visits under the commission regime ($6.86 < 8.50$, $p < .01$). These findings are in line with Results 3b and 3d in Table 2.

To probe these effects formally, we ran the following regression(s) separately for bonus and commission plans:

$$\log(A_{it}) = \alpha_0 + \sum_i \alpha_i S_i + \sum_t \gamma_t YM_t + \sum_m \upsilon^m \Gamma_{it}^m + \sum_m \kappa^m (\text{NewPlan})_t \times \Gamma_{it}^m + \delta_1 \log(\text{QtrTarget})_t + \delta_2 \text{GroupSize}_i + \varepsilon_{it}, \quad (14)$$

where A_{it} is the average of doctor (pharmacy) visits reported in territory i in period t and the remaining variables are all as defined previously. Note that although we include the vector Γ_{it}^m, which consists of the four dummy variables describing sales history in the last month of the quarter, we exclude the POST counterparts of these variables because we are not interested in timing effects here.

Table 9 reports the results from this analysis. We focus on doctor visits because the results for pharmacy visits are complementary. Recall that our predictions across the two plans are sharpest for salespeople who have achieved their quota. Specifically, Results 3b and 3d state that in the bonus plan, salespeople who have achieved/exceeded their quota will emphasize doctor visits, whereas those in the commission plan will emphasize pharmacy visits.

Our results support precisely this conclusion; the NewPlan coefficient is –.256, the EXCEEDED coefficient is +.288, and NewPlan × EXCEEDED is –.450, all significant

Table 8
MULTITASKING EFFECTS

A: By Regime		
	Bonus Plan	*Commissions Plan*
Doctor visits	10.08	7.93*
Pharmacy visits	4.92	6.74*

B: By Target Achievement	*Bonus Regime*		*Commissions Regime*	
Doctor Visits	*Before Reaching Quota (A)*	*After Reaching Quota (B)*	*Before Reaching Quota (C)*	*After Reaching Quota (D)*
Reached quota	8.02	12.98*	8.50	6.86*
Did not reach quota	10.10		7.91*	
Pharmacy Visits	*Before Reaching Quota (E)*	*After Reaching Quota (F)*	*Before Reaching Quota (G)*	*After Reaching Quota (H)*
Reached quota	6.3	3.37*	6.27	8.45*
Did not reach quota	4.90		6.71*	

*Denotes significant difference at .01 level.

Table 9
DOCTOR AND PHARMACY VISITS

Variables	*Pharmacy Visits*	*Doctor Visits*
NewPlan	.309*** (.005)	–.256*** (.004)
EXCEEDED	–.383*** (.022)	.288*** (.010)
NEAR	–.015 (.014)	–.009 (.008)
STRETCH	.007 (.008)	.005 (.003)
FAR	.002 (.013)	.0006 (.008)
NewPlan × EXCEEDED	.645*** (.024)	–.450*** (.021)
NewPlan × NEAR	.0386** (.016)	.0007 (.011)
NewPlan × STRETCH	.003 (.011)	–.0167 (.105)
NewPlan × FAR	.012 (.032)	–.0174 (.053)
GroupSize	.005 (.005)	.0026 (.003)
log(QtrTarget)	.011* (.006)	–.008 (.006)
Constant	1.51*** (.030)	2.329*** (.023)
Month–year effects	Included	Included
Fixed effects	Included	Included
Observations	7,107	7,392
Clusters	444	438
R-square (overall)	.4245	.3192

*$p < .1$.
**$p < .05$.
***$p < .01$.
Notes: Standard errors in parentheses, clustered by territory.

($p < .01$). Intuitively, in the bonus plan, the salesperson is willing to focus on doctor visits because there is no immediate opportunity cost. In the commission plan, however, salespeople continue to get commissions from any sales they book, so it makes sense for them to try to increase sales even after realizing their quota. None of the other results are statistically significant.

The finding that multitasking concerns are minimized after the bonus quota is reached is novel and makes an important theoretical contribution to the literature. It also provides a nuanced rationale for the existence of lump-sum bonuses, in that although the presence of a bonus provides an incentive for agents to increase effort, the ceiling imposed through a bonus enables the principal to engage agents in those activities that benefit the principal even though they are not directly compensated.

Overall, our results are largely supportive of our theoretical framework. Nevertheless, many alternative explanations remain, especially with respect to the overall identification in a before–after framework and issues related to turnover and ratcheting. We discuss these issues in the next section and present the details of some robustness checks in Web Appendix B (www.marketingpower.com/jmr_webappendix).

DISCUSSION

Detailed longitudinal data from a pre–post experimental design implemented at a single firm enabled us to uncover the relative merits of quota-bonus compensation plans versus quota-commission compensation plans. We summarize our empirical analysis into four key findings.

First, the switch to an equivalent commission plan improved sales productivity by 24% on average after accounting for industry, territory, and period effects. Second, the switch has heterogeneous effects at different ability levels. Specifically, the switch increases productivity at the lower deciles of ability much more so than at the higher deciles of ability. Third, sales groups engage in timing games under the bonus plan. Specifically, sales groups that either have achieved quota or are far from quota tend to reduce effort and push sales out to the next period; sales groups that can reach quota if they stretch tend to pull in sales from future periods. These effects are attenuated under the commission plan. Fourth, the bonus plan is much better than equivalent commissions in engaging salespeople to focus effort on tasks that are not directly related to short-term output (and thus incentive pay).

Assessing Threats to Validity

Our before–after design with multiple observations per territory over time rules out several validity threats arising from unobserved differences across territories by including effects for territory, time, and targets in our specifications. In Web Appendix B (www.marketingpower.com/jmr_webappendix), we describe some additional potential threats. Here, we briefly describe the supplemental analyses we undertook.

Individual versus multiperson territories. Recall that the firm only knows sales at the territory level, which includes both single-person and multiperson territories. We reran our regression models using only observations from single-person territories; the results were largely unchanged, suggesting no artifacts arising from multiperson territories.

Ratcheting. “Ratcheting” refers to quotas being raised in follow-on quarters when a salesperson has a good sales realization in a quarter. After learning this, salespeople may strategically adjust their effort downward. We raised this issue in our fieldwork and discovered that management was aware of this potential problem and assiduously worked to avoid it by excluding the immediate past quarter’s sales realization in the formula for the next quarter’s quota. Nevertheless, because past performance enters the quota-setting process along with company sales, industry sales, and territory features, we assessed ratcheting concerns directly using Misra and Nair’s (2011) specification of a reduced form regression that predicts quotas in period t from prior period sales and quotas. We found no empirical evidence of ratcheting (see Table B1 in the Web Appendix at www.marketingpower.com/jmr_webappendix).

Differential attrition. All experiments, including randomized trials, are threatened when subjects quit on account of their treatment. To this end, we compared quit rates before and after the intervention; the rates are virtually identical (see Table B2 in the Web Appendix). We also reran our regression models with additional measures accounting for turnover (see Table B3 in the Web Appendix).

Serially correlated outputs. Given the 36-month period of our observations, we reran our regression models to allow for serially correlated errors. We found no qualitative changes, ruling out unobserved time effects as a validity threat.

Despite these supplemental analyses, there are other limitations of the present experiment that warrant scrutiny in further research. First, we have not modeled our firm’s quota-setting process; a complete model would enable us to explicitly account for ratcheting concerns and other dynamic considerations. Second, our single-firm experiment raises generalizability issues; it would be worthwhile to study other institutional contexts.

Managerial Implications

Our results offer managerial implications for two common quota-based plans. For small sales task portfolios (which rule out large multitasking problems), commissions are preferable to lump-sum bonuses. Commission plans are also more desirable when a company aims for a smoother selling pattern (i.e., when timing games carry significant costs, commissions trump bonuses). In contrast, when firms need to incentivize their salespeople to produce more but the task portfolio is more diverse, bonus schemes more effectively encourage salespeople to engage in those tasks in the portfolio that do not have an immediate short-term impact on pay.

Our study joins other recent literature in marketing that focuses on understanding issues related to sales force productivity. Misra and Nair (2011) focus their energies on the issue of quota dynamics in general and the ratcheting effect in particular. Their prescription calls for implementing incentives without caps and using a shorter horizon for incentive payment. Notably, as we discussed previously, our firm seems to have overcome ratcheting concerns through the use of aggregate performance measures in updating quotas. As such, removal of caps is likely to lead to higher output, but our study suggests caution in implementing this prescription because it could potentially be counterproductive when significant multitasking concerns are present.

REFERENCES

Albers, Sönke and Murali Mantrala (2008), "Models for Sales Management Decisions," in *Handbook of Marketing Decision Models*, Berend Wierenga, ed. New York: Springer, 163–210.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2005), "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *The Quarterly Journal of Economics*, 120 (3), 917–62.

Chung, Doug, Thomas Steenburgh, and K. Sudhir (2010), "Do Bonuses Enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-Based Compensation Plans," Working Paper No. 1491283, Harvard Business School.

Gibbons, Robert (1987), "Piece-Rate Incentive Schemes," *Journal of Labor Economics*, 5 (4), 413–29.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan (2003), "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 111 (3), 465–97.

Hardin, Garrett (1968), "The Tragedy of the Commons," *Science*, 162 (3859), 1243–48.

Holmstrom, Bengt and Paul Milgrom (1991), "Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, & Organization*, 7 (Special Issue), 24–52.

Jain, Sanjay (2009), "Self-Control and Optimal Goals: A Theoretical Analysis," *Marketing Science*, 28 (6), 1027–45.

Jensen, Michael C. (2003), "Paying People to Lie: the Truth About the Budgeting Process," *European Financial Management*, 9 (3), 379–406.

Joseph, Kissan and Manohar U. Kalwani (1998), "The Role of Bonus Pay in Salesforce Compensation Plans," *Industrial Marketing Management*, 27 (2), 147–59.

Knez, Marc and Duncan Simester (2001), "Firm-Wide Incentives and Mutual Monitoring at Continental Airlines," *Journal of Labor Economics*, 19 (4), 743–72.

Latham, Gary P. and Edwin A. Locke (1991), "Self-Regulation Through Goal Setting," *Organizational Behavior and Human Decision Processes*, 50 (2), 212–47.

Lazear, Edward P. (1995), *Personnel Economics (Wicksell Lectures)*. Cambridge, MA: MIT Press.

——— (2000), "Performance Pay and Productivity," *American Economic Review*, 90 (5), 1346–61.

Misra, Sanjog and Harikesh S. Nair (2011), "A Structural Model of Sales-Force Compensation Dynamics: Estimation and Field Implementation," *Quantitative Marketing and Economics*, 9 (3), 211–57.

Oyer, Paul (1995), "The Effect of Sales Incentives on Business Seasonality," Working Paper No. 354, Industrial Relations Section, Princeton University.

——— (1998), "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality," *The Quarterly Journal of Economics*, 113 (1), 149–85.

——— (2000), "A Theory of Sales Quotas with Limited Liability and Rent Sharing," *Journal of Labor Economics*, 18 (3), 405–426.

Paarsch, Harry J. and Bruce Shearer (2000), "Piece Rates, Fixed Wages, and Incentive Effects: Statistical Evidence from Payroll Records," *International Economic Review*, 41 (1), 59–92.

Prendergast, Canice (1999), "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37 (1), 7–63.

Steenburgh, Thomas J. (2008), "Effort or Timing: The Effect of Lump-Sum Bonuses," *Quantitative Marketing and Economics*, 6 (3), 235–56.

Zoltners, Andris A., Prabhakant Sinha, and Sally E. Lorimer (2008), "Sales Force Effectiveness: A Framework for Researchers and Practitioners," *Journal of Personal Selling and Sales Management*, 28 (2), 115–31.