The Effects of Quota Frequency: Sales Performance and Product Focus

Doug J. Chung
Das Narayandas
Dongkyu Chang

Working Paper 17-059
The Effects of Quota Frequency: Sales Performance and Product Focus

Doug J. Chung
Harvard Business School

Das Narayandas
Harvard Business School

Dongkyu Chang
City University of Hong Kong

Working Paper 17-059
The Effects of Quota Frequency: Sales Performance and Product Focus

Doug J. Chung, Harvard University
Das Narayandas, Harvard University
Dongkyu Chang, City University of Hong Kong*

August 2018

Abstract

This study investigates the comprehensive and multidimensional effects of quota frequency on sales force performance. We develop a theory of salespeople’s behavior with regard to the effect of sales quota frequency on aggregate effort and the product type focus. The theory includes many realistic elements such as salespeople’s multi-dimensional effort, heterogeneity in ability, product focus, and forward-looking behavior. We test our theory through a field experiment in which we vary the sales force compensation scheme of a major retail chain in Sweden. We find that shifting to a temporally more frequent quota plan leads to an increase in sales performance for low-performing salespeople by preventing them from giving up in the later periods within a quota evaluation cycle. However, we find little evidence of improvement in productivity for high-performing salespeople. In addition, we find no effects in product returns with regard to a change in quota frequency. With quotas set over shorter time horizons, the high-performing salespeople focus mainly on low-ticket products, resulting in a decrease in both sales volume and the sale of high-ticket products, thus reducing the firm’s profits.

Key words: sales force compensation, field experiment, quota, goal, product focus.

* Doug J. Chung is the MBA Class of 1962 Associate Professor of Business Administration (dchung@hbs.edu) and Das Narayandas is the Edsel Bryant Ford Professor of Business Administration (nnarayandas@hbs.edu) at the Harvard University. Dongkyu Chang is an Assistant Professor of Economics at the City University of Hong Kong. The authors thank the seminar participants at the California State University, Harvard University, Sungkyunkwan University as well as the conference participants at the 2017 Consortium for Operational Excellence in Retailing (COER), 2017 INFORMS Marketing Science, and the 2018 Yale Customer Insights conferences for their comments and suggestions. The authors also thank the anonymous firm for cooperating in this study.
1. Introduction

Incentives are ubiquitous, especially in a capitalistic free-market economy. They are believed to provide one of the primary motivations for people to work, especially in the domain of personal selling. Despite recent advances in sophisticated marketing techniques using big data to persuade customers and encourage purchases, personal selling still remains a significant (and in most industries the only) function in firm–customer interactions. According to the U.S. Bureau of Labor Statistics, 14 million people, or about 10% of the entire U.S. labor force, are employed in personal selling (U.S. Department of Labor, 2012). U.S. firms spend over $800 billion annually on personal selling, an amount that is 4.7 times greater than the total expenditures on all media advertising ($169.5 billion) and more than 20 times greater than total spending on Internet advertising ($39.5 billion) (Zoltners, Sinha, & Lorimer, 2013). A large proportion of spending on personal selling is used to incentivize and motivate salespeople to exert greater selling effort. The sales force compensation plan is considered to be the managers’ primary tool for incentivizing and motivating salespeople. With so many people and resources at stake, the design of the sales force compensation plan is of great strategic importance to firms.

A sales force compensation plan typically consists of a fixed salary plus variable pay conditional on meeting a sales quota (i.e., achieving a certain threshold of performance). Firms commonly use quotas; about three quarters of U.S firms use some form of quotas (Joseph & Kalwani, 1998). Figure 1 shows illustrative examples of several quota-based plans. Firms typically use quotas as achievement goals to evaluate performance and determine whether a salesperson has had a successful period (e.g., month or year). But how should a sales manager
design a quota-based compensation plan? In this research, we specifically attempt to answer, through theoretical illustration and empirical evidence via a field experiment, the following questions: What is the appropriate frequency of quotas? That is, at what intervals should quotas be set, and how often should they be evaluated? Would frequent quotas either increase or decrease sales performance? \(^1\) If so, which types of salespeople would be affected? Does a salesperson’s quality of effort falter with frequent quotas? Would there be a change in a salesperson’s product focus if the frequency of sales quotas were changed?

To answer these questions, we first develop a theoretical model of salespeople's behavior with regard to different quota frequency. Then, we conduct a field experiment with full-time salespeople to validate our theoretical predictions. In our theoretical model, we postulate that salespeople’s actions are multi-dimensional and that their selling abilities differ with regard to the product type (low-ticket versus high-ticket). We provide a general proof that it is in the best interest of the firm to provide the most frequent quota evaluation cycles for salespeople who have limited ability to sell high-ticket products. In contrast, it may be optimal for the firm to have less frequent quota cycles for salespeople who are better at selling high-ticket products. The firm wants salespeople to focus their efforts on products that they are effective at selling—that is, salespeople with the ability to sell high-ticket products should focus on selling such products. However, because the incentives of the firm and the salespeople are misaligned (the firm does not incur the salespeople’s disutility of effort), there can be a distortion in the salespeople’s actions. As a consequence, with

---

\(^1\) We refer to “frequent quotas,” “frequent quota plans” or “short quota cycles” as quota-based plans that have more evaluation and payment periods than less frequent quota plans. In our empirical context, a daily quota plan represents a frequent quota plan, and a monthly quota plan represents a less frequent quota plan.
frequent quotas, salespeople that are a better fit for selling high-ticket products may alter their behavior to focus on low-ticket products, changing the type of products sold and, thus, decreasing the firm’s aggregate sales.

To validate our theoretical findings and empirically examine the effects of quota frequency on sales performance, we collaborated with a major retail chain in Sweden to conduct a field experiment with full-time sales employees. We implemented an intervention in which, holding everything else constant, we changed the sales force compensation scheme from a monthly to a daily quota plan. We also deployed a control group of several stores whose salespeople did not encounter a change in compensation structure during the observation period. We use the variation in performance between the salespeople who experienced the change and those who did not to account for any seasonal and other exogenous fluctuations in order to analyze, as cleanly as possible, the pure causal effect of quota frequency on various dimensions of sales performance.

Overall, in this paper, we uniquely contribute to the literature in several ways. First, we develop a theoretical model of the effect of quota frequency on salespeople’s behavior. The model includes many realistic elements such as multi-dimensional effort (with regard to the product focus) and forward-looking behavior; combined, these predict not only aggregate sales outcomes but also changes in the types of products sold. Second, we conduct a large-scale field experiment involving more than 300 full-time salespeople of a major retail chain. We do so by exogenously changing the compensation plan along with a control group to account for any normal/temporal changes so that we can isolate the causal effect of quota frequency on sales performance. Third, to the best of our knowledge, this study is one of the first to directly examine quota frequency in a sales force
setting. Fourth, we empirically investigate the effect of quota frequency not only on sales revenue, but also on various other dimensions of performance, such as the quality and type of effort, by monitoring and measuring product returns and changes in product focus. Finally, we examine the heterogeneous effects of quota frequency across salespeople of different types.

Substantively, through theoretical illustration and the empirical results of our field experiment, we find that a change from a less frequent (monthly) to a more frequent (daily) quota plan increases sales performance, but mainly for low-performing salespeople. Because every day is a fresh start under the daily plan, salespeople’s motivations are intact throughout the month, whereas under the monthly plan, salespeople give up if they are too far from meeting quota in the later days of the month. In contrast, we find potentially negative effects of quota frequency on high-performing salespeople, as they distort their efforts from high-ticket to low-ticket products, resulting in a change in the types of products sold and a decrease in aggregate sales, thus reducing the firm’s profits. Surprisingly, in the daily quota plan, we do not find any evidence of overaggressive selling that results in an increase in product returns, even though salespeople were not penalized for returns.

The rest of the paper is organized as follows. In Section 2, we summarize the related literature, and in Section 3, we present our theoretical model. In Section 4, we explain the institutional details of the firm and the field experiment design. Section 5 presents the empirical model-specification, and Section 6 discusses the results. Section 7 concludes.

---

2 Chung, Steenburgh, and Sudhir (2014) explored the concept of quota frequency. However, their analysis was based on counterfactual simulations using estimates from their structural model and not inferred directly from the data.

3 We define high-performing salespeople as those who are effective at selling high-ticket products.
2. Literature Review

Despite the ubiquitous use of sales quotas, academics have remained skeptical about their effectiveness. In the economics literature, assuming rational agents, there are two primary arguments against the use of sales quotas. First, the discrete and nonlinear nature of quotas commonly pushes salespeople to less powering areas of incentives (Holmstrom & Milgrom, 1987; Lal & Srinivasan, 1993). That is, the motivational effects of achieving a sales quota set by the firm diminish when a salesperson either has already surpassed the quota or is too far from achieving it—a salesperson who has no hope of meeting quota is likely to give up. Second, in a B2B environment (less so in a B2C retail environment such as our empirical setting), sales quotas may only provoke salespeople to manipulate the timing of sales such that they have no additional effect on performance (Oyer, 1998). For example, if a salesperson has already met quota, instead of booking realized sales in the current period, he or she can simply ‘push’ sales into the future to count toward quota in the next period. Relatedly, a salesperson who has not met quota can ‘pull’ sales from the future (i.e., book scheduled future sales in the current period) to attain quota. Increasing the temporal frequency of quotas (e.g., from monthly to daily) would make a quota-based plan similar to a linear commission plan. Hence, according to the above two arguments, frequent quotas would provide better constant motivation to salespeople regardless of past cumulative sales and would mitigate the timing manipulation of sales.

There is a vast literature in psychology on goals—such as sales quotas—and their effect on motivation (for an extensive survey, see Latham & Locke (1991)). We discuss this literature in two parts—that of goals, in general, and that of subgoals.
The “goal-gradient hypothesis” (Hull, 1932; Hull, 1938) postulates that people become more motivated—the goal gradient gets steeper—as their (perceived) progress gets closer to a goal (Cheema & Bagchi, 2011; Kivetz, Urminsky, & Zheng, 2006; Nunes & Drèze, 2006). A goal, by definition, is a key reference point with regard to a focused activity (Heath, Larrick, & Wu, 1999). Hence, the characteristics of the value of attaining a goal can be similar to those of the value function in Prospect Theory (Kahneman & Tversky, 1979)—losses loom larger than gains, diminishing sensitivity from the origin: concave in gains and convex in losses. As a result, as a person gets closer to a goal, his or her marginal motivation to attain the goal becomes higher (steeper goal gradient).

Because sensitivity diminishes as one is further from a goal, a person would be less motivated in the initial stages; this is known as the ‘starting problem’ (Heath, Larrick, & Wu, 1999). Splitting a grand goal into multiple smaller subgoals will move a person relatively closer to the origin (reference point), mitigating the starting problem (i.e., sensitivity and, thus, motivation increases). However, there can also be a negative motivational effect of multiple subgoals compared to one grand goal. Although motivation increases as ones gets closer to a subgoal, it will substantially decrease (flat goal gradient) once the subgoal is achieved. That is, people will become complacent and reduce effort after achieving subgoals (Heath, Larrick, & Wu, 1999). Furthermore, attainment of subgoals may liberate or permit a person to pursue other goals (Amir & Ariely, 2008; Fishback & Dhar, 2005; Fishbach, Dhar & Zhang, 2006).

Although there are both positive and negative motivational effects with regard to attaining subgoals, past research has found that the positive effects outweigh the negative effects—for
example, Gal & McShane (2012) in the domain of debt management and Zhang & Gao (2016) in the domain of reward programs. Relatedly, in the education literature, researchers have found that frequent testing results in better performance outcomes for students (for an extensive survey, see Bangert-Drowns, Kulik, & Kulik (1991)).

The above survey of relevant literature suggests that, in our empirical context, having more frequent quotas (by splitting a monthly quota into many smaller daily quotas) should lead to greater motivation and, thus, higher sales performance. There are two primary reasons. First, because a daily quota plan gives salespeople a fresh start each day, it should help them maintain high motivation throughout the month. For example, under a monthly quota plan, a salesperson who had a series of bad luck early in the month may decide to give up late in the month because there is no chance of meeting or exceeding the firm’s quota. This would not be the case under the daily quota plan, as every day would present the salesperson with a new chance to succeed. Second, the daily quota plan would tap into a salesperson’s motivation more often, and, thus, there would be more instances of steeper goal-gradient.

However, as Darmon (1997) indicates, to motivate salespeople to achieve objectives, quotas should be challenging. Splitting a grand quota into multiple finer (thus more frequent) quotas would be the same as replacing one challenging quota with many less challenging quotas. Using the principal-agent theory framework, Park (1995), Kim (1997) and Oyer (2000) illustrate that, under specific assumptions, a non-linear discrete quota-bonus plan (i.e., longer quota cycles in our domain) can be optimal for the firm. Furthermore, the flexibility to intertemporally allocate effort across multiple periods may make the monthly plan more effective, as a change to a daily quota
plan may merely provoke income targeting within a day (Camerer et al., 1997). In addition to the abovementioned arguments, the daily quota plan could potentially increase anxiety and stress among salespeople, as they may constantly worry about meeting quota day in and day out, resulting in demotivation. Also, as we witnessed with Sears in the 1990s, Marsh in the 2000s, and more recently with Wells Fargo, there can also be negative effects of an overly aggressive incentive compensation plan (Zoltners, Lorimer, & Sinha, 2016), leading to unethical behavior and fraud (Schweitzer, Ordóñez, & Douma, 2004). In our context, the daily quota plan may induce salespeople to become overly aggressive by selling products that would ultimately be returned later, hurting the firm in the long run.

We have, thus far, discussed both the positive and negative effects of a frequent quota compensation plan on aggregate sales, assuming unidimensional effort. However, a salesperson’s effort is naturally multidimensional. Hence, the compensation plan, in addition to affecting aggregate sales, may alter various dimensions of effort (Holmstrom & Milgrom, 1991). For example, Kishore et al. (2013) find that a switch from a discrete bonus to a commission plan increases effort on incentivized tasks but decreased effort on non-incentivized tasks. Thus, a change in compensation may increase aggregate sales but, at the same time, adversely affect other dimensions of performance. We seek to gain insights by examining changes in salespeople’s product focus with regard to the quota frequency of the compensation plan.

In summary, it is unclear how a change in the quota frequency of the compensation scheme would affect sales performance across multiple dimensions. We develop, to the best of our knowledge, the first theoretical model of salespeople’s behavior with regard to product focus under
different quota compensation plans. Furthermore, to validate our theoretical claims, we conduct a field experiment with full-time sales employees evaluated and compensated under different quota frequencies, which again, to the best of our knowledge, is the first. We present our theoretical model and empirical applications in the following sections.

3. Theoretical Model of an Agent’s Behavior based on Quota (Evaluation) Frequency

3.1. Model Setup

We consider a job of a salesperson\(^4\) (agent) who can choose to focus his or her selling efforts on either a low-ticket product \((L)\) or a high-ticket product \((H)\) in periods \(k = 1, 2, \ldots\). If the agent focuses on \(\theta \in \{L, H\}\) in period \(k\), sales volume \(x_k = \chi^\theta\) is generated with a probability \(0 \leq p^\theta < 1\), and no sale occurs with the complementary probability. The agent can also take the null action \((\emptyset)\) where the probability of sales occurring is zero. We assume that \(\chi^H > \chi^L > 0\).

The agent receives commission-based compensation at the end of each quota cycle. A quota cycle consists of \(N > 0\) discrete periods \(k = 1, 2, \ldots, N\), and the agent’s commission rate for each quota cycle depends on the per-period average sales volume within the quota cycle.\(^5\) Formally, let \(X = \sum_{k=1}^{N} x_k\) denote the total sales volume generated within a quota cycle. The agent’s commission rate for this quota cycle is \(\beta(X / N) \geq 0\), and, thus, the total commission that the

\(^4\) Hereafter, we use the term salesperson and (sales) agent interchangeably to refer to an employee hired by the firm to engage in its sales activities.

\(^5\) This setting can easily be extended to a quota structure consisting of total sales volume.
agent receives will be $\sigma_N(X) = \beta(X / N)X$. We assume that $\beta$ is non-decreasing and piecewise-continuous. We also assume that $\beta(\chi^H) > 0$ without loss.\(^6\)

The agent’s total payoff in each quota cycle is the sum of the total compensation and the disutility (cost) from exerting effort. Let $a_k \in \{a^H, a^L, a^0\}$ denote the agent’s action in period $k$, where $a^H$ and $a^L$ indicate the agent’s actions focusing on $H$ and $L$ products, respectively, and $a^0$ indicates the null action. The disutility from each action is given by $c(a^H) = c^H > 0$, $c(a^L) = c^L > 0$, and $c(a^0) = 0$, where we assume that $c^H \geq c^L > 0$. The agent maximizes the expected total payoff

$$E\left[\sigma_N(X) - \sum_{k=1}^{N} c(a_k)\right].$$

A policy of the agent specifies which action to take in each period, contingent on the agent’s actions and realized sales in the previous periods within a quota cycle. Let $\Sigma_N$ denote the set of all agent-optimal policies that maximizes the expected total payoff. We assume that the firm wants to maximize the (per-period) expected revenue. Because the firm does not incur an agent’s disutility of effort, a distortion may exist, and, thus, the agent-optimal policy may not be optimal for the firm. Let

$$\mu_N = \sup\left\{E'\left[\frac{X}{N}\right] : \sigma \in \Sigma_N\right\}$$

\(^6\) $X / N$ never exceeds $\chi^H$, and, thus, $\beta(X / N) \leq \beta(\chi^H)$ with probability 1. Therefore, if $\beta(\chi^H) = 0$, the commission is always zero regardless of the agent’s performance, and, thus, the agent never exerts effort.
denote the expected per-period revenue as a function of $N$, where $\mathbb{E}^{\sigma}[X/N]$ is the expected average sales volume under an agent-optimal policy $\sigma$. Note that both $\Sigma_N$ and $\mu_N$ depend on $N$.

### 3.2. Agent’s Choice of Effort in the Product Focus

We can decompose the agent’s decision into two parts: 1) the quantity and 2) the class of effort. First, the agent chooses whether or not to exert effort, which we define as the quantity of effort. Second, conditional on exerting effort, the agent also chooses which product type to focus on (either high-ticket or low-ticket products) which we define as the class of effort.

The firm cares about both the quantity and the class of effort. Note that the firm may not want the agent to focus solely on high-ticket products. First, if $p^H \chi^H < p^L \chi^L$, the firm prefers the agent to exert effort on low-ticket products because the expected revenue is higher for low-ticket than for high-ticket products. Second, even if $p^H \chi^H > p^L \chi^L$, the firm may not want the agent to divert effort to high-ticket products if such a change is accompanied by a significant reduction in the quantity of the agent’s effort. Simply put, the firm, in each period, wants agents to exert effort versus no effort. In addition, the firm wants agents to exert effort consistent with their types—i.e., agents who are effective at selling high-ticket products should focus on those products, and vice versa.

### 3.3. Agent Heterogeneity
We focus on the comparative static analysis with respect to $p^H$, the parameter we interpret as the agent’s ability to promote high-ticket products. High-ticket products typically are sophisticated and technologically advanced, needing substantial information for usage and handling. Sales agents usually help to provide such information to potential consumers. Moreover, many new high-ticket products are experience goods, and consumers often rely on sales agents’ assessment of their value before purchase. Hence, consumers require more sales assistance when considering the purchase of high-ticket products than when considering low-ticket products. The quality of such sales assistance is highly dependent on a sales agent’s ability, which typically is highly heterogeneous. In contrast, low-ticket products need less sales assistance, and the ability to sell them is relatively homogeneous across sales agents. In what follows, we refer to a person with a high value of $p^H$ as a ‘high-type’ sales agent and a person with a low value of $p^H$ as a ‘low-type’ sales agent.

3.4. Optimal Quota Frequency for Low-Type Agents

We first consider the provision of incentive for a low-type sales agent. The following proposition shows that it is optimal for the firm to set the most frequent quota cycle ($N = 1$) for a low-type agent.

**Proposition 1.** There is $p^H > 0$ such that, whenever $p^H \in [0, p^H]$, $\mu_1 \geq \mu_N$ for all integer $N > 1$.

The formal proof of the above proposition is stated in the Appendix. The strength of this proposition stems from its generality and robustness. The proposition imposes no restriction on
the compensation scheme (i.e., the commission rate $\beta$) beyond monotonicity and piecewise-continuity, which are observed in virtually all commission schemes used in practice. Also, note that the proposition assumes that the agent is fully rational and forward-looking. This stands in contrast to other existing theories that explain the effectiveness of frequent goals (quota cycles) based on psychological/behavioral assumptions (Hull, 1932; Hull, 1938; Heath, Larrick, & Wu, 1999; Kivetz, Urmsisky, & Zheng, 2006; Nunes & Drèze, 2006; Cheema & Bagchi, 2011; Gal & McShane, 2012; Zhang & Gao, 2016).

The intuition for this proposition is relatively straightforward. When $p^H$ is close to zero, under any quota cycle, it is always suboptimal for the agent to exert effort on high-ticket products. Hence, the firm is concerned only with maximizing the total quantity of the agent’s effort, without worrying about the class of effort. In other words, it is in the firm’s best interest to make the agent exert effort on low-ticket products (instead of the null action) as much as possible.

Relatedly, in terms of intertemporal dynamics, if $N > 1$, it occurs with positive probability that the agent’s effort fails to realize in the early periods of the quota cycle. In this case, the agent becomes hopeless in meeting quota and obtaining high compensation at the end of the quota cycle. Consequently, the agent will give up and stop working in later periods. If $N = 1$, on the other hand, bad luck in early periods does not affect the agent’s incentive in later periods, and, thus, the agent exerts effort in all periods. Hence, the shortest quota cycle ($N = 1$) is the most effective plan for incentivizing low-type sales agents.

Let us illustrate an agent’s optimal policy with the quota cycle $N = 1$. Suppose that the agent exerts effort on high-ticket products. The agent suffers from disutility $e^H$, regardless of the sales
outcome, which will realize to $\chi^H$ with probability $p^H$ or to zero with probability $1 - p^H$. Hence, the agent’s expected utility from exerting effort on high-ticket products is

$$u^H = p^H \beta(\chi^H)\chi^H - c^H.$$  

Similarly, the expected utility from exerting effort on low-ticket products is

$$u^L = p^L \beta(\chi^L)\chi^L - c^L.$$  

We can see that $u^H \geq u^L$ if and only if

$$p^H \geq A(p^L) \equiv \frac{\beta(\chi^L)\chi^L}{\beta(\chi^H)\chi^H} p^L + \frac{c^H - c^L}{\beta(\chi^H)\chi^H}.$$  

(1)

**Figure 2a** describes the optimal action that maximizes the agent’s payoff for each combination of $(p^L, p^H)$. The agent focuses on high-ticket products if his or her $(p^L, p^H)$ lies in the horizontally shaded area, and on low-ticket products if his or her $(p^L, p^H)$ lies in the gridded area. The agent will take the null action in all other cases because both $u^H$ and $u^L$ are negative. Simply put, the agent exerts effort on the product types that are consistent with his or her relative ability if the minimum expected utility when exerting effort is satisfied (greater than zero).

**Figure 2b** is identical to **Figure 2a**, except for the line initiated from the origin. This line is the collection of points $(p^L, p^H)$ at which both high- and low-ticket products generate the same expected revenue for the firm. If $(p^L, p^H)$ lies in the area above (below) the line, the firm prefers the agent to focus on high-ticket (low-ticket) products. The gridded area in **Figure 2b** represents
the case that is the optimal quota frequency: the agent focuses on low-ticket products to maximize his or her expected utility, which happens to also maximize the firm’s expected revenue.\(^7\)

3.5. Optimal Quota Frequency for High-Type Agents

We showed in Proposition 1 that setting a frequent quota cycle \((N = 1)\) is effective in motivating low-type agents, maximizing the quantity of their effort. However, setting a frequent quota cycle may not be effective in incentivizing high-type agents. In fact, the effect of quota frequency on high-type agents is much more compound and subtler than its effect on low-type agents.

See the diagonally shaded triangular area in Figure 2b. It is optimal for the firm for agents in this area to focus on high-ticket products; yet they focus on low-ticket products under the quota cycle \(N = 1\). Hence, the firm may utilize a longer quota cycle compensation scheme to divert agents’ efforts from low-ticket to high-ticket products. However, the firm has to be cautious. An agent may change the class of effort, as desired by the firm but, at the same time, may also reduce the quantity of effort under a longer quota cycle, especially after bad luck in the early stages (because the agent is unlikely to meet quota no matter how hard he or she works in the remaining periods of the quota cycle). If this is the case, diverting the agent’s effort to high-ticket products is not necessarily beneficial to the firm.

\(^7\) The gridded area includes the segment \(((p^l, p^h): 0 < p^l < c^l / \beta (X^l) \lambda^l \text{ and } p^h = 0)\) on the horizontal axis. These agent types take the null action under any quota frequency, and, thus, the quota frequency of \(N=1\) is vacuously optimal for the firm.
The following example shows that setting a longer quota cycle indeed has both positive and negative effects on the incentives of high-type agents.

**Example 1.** Consider the following commission scheme: commission rate $b > 0$ is applied if the average sales volume is larger than $y > 0$, and otherwise, zero. That is,

$$
\beta(X / N) = \begin{cases} 
    b & \text{if } X / N \geq y \\
    0 & \text{otherwise.}
\end{cases}
$$

We assume a reasonably attainable quota $\chi^L / 2 < \chi^H / 2 < y \leq \chi^L$. We also assume that $c^H / \chi^H < b < (c^H - c^L) / (\chi^H - \chi^L)$.  

First, consider the case $N = 1$. The two lines $p^H = A(p^L)$ and $p^H \chi^H = p^L \chi^L$ are depicted in [Figure 3a](#), where $A(p^L)$ is as defined in Equation (1). If $N = 1$, as discussed in the previous subsection, agents in the gridded area focus on low-ticket products, even though it is more profitable for the firm if they focus on high-ticket products. That is, the class of the agent’s effort is not satisfactory from the firm’s perspective. However, the quantity of the agent’s effort is satisfactory because the agent never takes the null action.

Now suppose that the firm changes its quota frequency to $N = 2$. We can show that all agents in the gridded area reduce their quantity of effort after this change. To see this, focus on the case in which an agent earned no revenue in the first period. Because $\chi^L / 2 < \chi^H / 2 < y$,

---

8 Setting an upper bound for the commission rate $b$ is consistent with our empirical setting of commission at quota with low commission rates (in the range between 0.27% and 2.0%). Setting the lower bound $c^H / \chi^H$ excludes the less interesting case in which agents never exert effort in high-ticket products.

9 The two lines have the same slope due to our assumption $y \leq \chi^H$ (reasonably attainable quota).

10 If both $p^L$ and $p^H$ are strictly less than one (as we have assumed), zero revenue occurs in the first period with a positive probability under any policy of the agent.
the agent can never achieve the quota $X / N \geq y$, no matter how much effort he or she exerts in the next period. Hence, the agent would give up and take the null action in the second period. This stands in contrast to the case $N = 1$, in which the agent never takes the null action, and, hence, it shows that the quantity of effort decreases.

On the other hand, by changing the quota frequency to $N = 2$, the firm can improve on the class of effort for a subset of these agents. Consider the case in which an agent earned positive revenue $X_i \in \{\chi^H, \chi^L\}$ in the first period. If the agent focuses on a high-ticket product in the second period, the agent’s expected continuation payoff would be $v^H(X_i) \equiv p^H b(X_i + \chi^H) - c^H$. Similarly, the expected continuation payoff from focusing on a low-ticket product is $v^L(X_i) \equiv p^L b(X_i + \chi^L) - c^L$. The agent chooses $a_i = a^H$ if and only if $v^H(X_i) > v^L(X_i)$, and, thus,

$$p^H > B(p^L; X_i) = \frac{X_i + \chi^L}{X_i + \chi^H} p^L + \frac{c^H - c^L}{(X_i + \chi^H)b}.$$ 

One can easily check that $A(p^L) > B(p^L; \chi^L) > B(p^L; \chi^H)$ for any $0 \leq p^L < 1$ under our assumption $b < (c^H - c^L) / (\chi^H - \chi^L)$. Therefore, agents in the dotted area in Figure 3b exert effort on high-ticket products in the second period, conditional on $X_i > 0$.

The intuition for diverting effort to high-ticket products is as follows. First, note that agents in the dotted area are more productive in promoting high-ticket products ($p^L \chi^L < p^H \chi^H$).

---

11 Non-zero revenue is generated in the first period with positive probability under the agent-optimal policy if, for example, $c^L/p^L$ is sufficiently small. In this case, the agent’s expected benefit from exerting effort on a low-ticket product exceeds the disutility $c^L$; hence, the agent never takes the null action in the first period.

12 If the agent exerts effort on high-ticket products in the second period, the expected revenue is $\chi^H$, and, thus, the final compensation is $b(X_i + \chi^H)$ with probability $p^H$. With the complementary probability, the revenue in the second period is zero, and, thus, the final compensation is zero (because the resulting average sales volume $X_i/2 < y$ cannot achieve the quota). The agent takes into account only the disutility incurred in the second period, disregarding the disutility already incurred in the first period.
However, these agents would rather focus on the less productive action when $N = 1$ because the expected compensation from focusing on a high-ticket product, $p^u(b\chi^u)$, does not sufficiently cover their disutility $c^u$. Next, consider the agent’s problem in the second period, conditional on $N = 2$ and $X_1 > 0$. The disutility from high-ticket products remains the same. However, the expected reward is now $p^u b(X_1 + \chi^u)$ and strictly larger than the compensation $p^u(b\chi^u)$ under $N = 1$. Hence, agents become more willing to exert effort on a high-ticket product after the change in the quota frequency. Note that the already-cumulated revenue $X_1$ diverts agents to a more productive action.\footnote{This intuition is closely related to the goal-gradient hypothesis which also predicts that an agent is willing to exert more effort as he or she gets closer to a goal. Note that in our case, however, agents are assumed to be fully rational and forward-looking. Therefore, the goal-gradient-hypothesis-like behavior in this example is driven by the structure of the compensation scheme rather than by the behavioral characteristics of the agents. Also, note that our observation here concerns the class of the agent’s effort, while the goal-gradient hypothesis focuses mostly on the quantity of effort.}

In summary, there is a trade-off between the agent’s class and quantity of effort when the firm changes its quota frequency. Changing to a longer quota cycle of $N = 2$ reduces the quantity of the agent’s effort but improves the class of the agent’s effort. Such a trade-off is not specific to this example. Utilizing a longer quota cycle, the firm may improve in the dimension of the class of the agent’s effort by diverting the agent to a more productive task, but the agent may reduce the quantity of effort. Which effect dominates the other is an empirical question, which we seek to answer in the subsequent sections.

When heterogeneous agents are simultaneously employed, the firm may offer a menu of compensation schemes and allow agents of different types to self-select into different commission schemes, as suggested by the literature on mechanism design (Laffont & Martimort, 2002, Chapter
2). However, such a complicated compensation system can be practically difficult to implement. Also, different commission schemes across agents may cause a concern over fairness and, thus, decrease overall morale. As a result, changing a component (such as a quota cycle) of the compensation scheme uniformly across all agents is often the only feasible option for the firm. As illustrated above, such a change likely has different effects across different types of people, and, thus, the firm has to be cautious and aware of the expected aggregate outcome as a result of a change in the compensation structure.

4. Institutional Details and the Field Experiment Design

To validate our theoretical claims in the previous section, we conducted a field experiment with full-time sales employees, varying the compensation plan’s quota frequency. The focal firm is a highly regarded retail chain operating 94 stores in Sweden. It sells mostly accessories for cellular phones and home electronics (e.g., networking accessories, headsets and phone cases) and parts for consumer electronics and home appliances (e.g., semiconductors and switches). It also sells small-to-medium-sized consumer electronic goods, such as data-storage devices, network appliances, DVD players, and wireless routers. Product prices range from less than $1 to $500 or more, with an average price of slightly over $20. All of the stores are company-owned, and the firm employs a direct-sales force of about 350 salespeople at any given time across its stores. The compensation plan for sales employees consists of a fixed salary plus a variable commission on sales. The commission rate (and, thus, the commission amount) is determined by sales performance, measured in average sales per hour (SPH).
The details of the variable component of the compensation plan are given in Table 1. There are five levels of commission a salesperson can receive. For example, if a salesperson’s average SPH were $150 at the end of the evaluation period, he or she would receive a commission of 0.27% for every dollar of sales. If a salesperson’s average SPH were $250 or more, he or she would receive the highest commission level of 2.0%. Note that the quotas are in average SPH instead of in absolute amounts. Because of this characteristic, along with the discrete nature of the tiered commission levels, a salesperson’s variable pay would have a kink at each tier (quota) level, which resembles a quota-based commission plus lump-sum bonus scheme (Figure 1d). Figure 4a illustrates the level of variable pay for a salesperson who is assigned 140 hours a month. The figure shows that, as a salesperson achieves each quota level, there is a step jump in pay due to discretely accelerating commission rates. A salesperson would make $1,000 in variable pay if his or her monthly sales totaled $50,000.

The field experiment and, thus, the change in the sales force compensation plan took effect on May 1, 2015. Holding everything else constant, including the commission rate per quota achieved and the quotas in terms of average SPH (Table 1), only the evaluation period changed (from a month to a day) on May 1. That is, up until the end of April 2015, the commission rate for each salesperson was assessed by summing all sales that person made and dividing them by all hours that person worked within a month. Then, starting on May 1, the commission rate was evaluated daily. Figure 4b illustrates commission pay as a function of daily sales for a salesperson who is assigned seven hours a day. The overall shapes of Figure 4a and Figure 4b look very similar, as
only the frequency of evaluation changed, while the commission rate and the quotas in Table 1 remained the same.

In addition to changing the compensation structure for one group of employees across the firm’s stores (the treatment group), we also arranged for several stores and their salespeople to experience no changes in the compensation plan during the experiment (the control group). Hence, we use the difference in sales performance between the treatment and the control group to identify the magnitude of the treatment effect (the daily quota plan), taking into account any normal/temporal changes (e.g., seasonality or firm-level advertising) in sales that would have occurred regardless of the change in the compensation plan.

With help from management, we chose stores for the control group that embodied a representative sample of stores across geographical areas of the country.¹⁴ The majority of Sweden’s population resides in the southern tip of the country, concentrated in the suburbs and city centers of the three most populous cities, Stockholm, Gothenburg, and Malmö. We confronted two main challenges in choosing our control stores.

First, to avoid complications in implementing the changed compensation plan, the focal firm initially did not want any control stores. Also, the firm’s management was extremely concerned about fairness across employees. Many members of management (including the sales director, the information technology manager, and the vice president of operations) had risen through the ranks, starting out as in-store salespeople, and fairness was one the firm’s primary human resources (HR)

---

¹⁴ The firm had only eight stores in the entire central and northern part of Sweden; we omitted these stores from our analysis, leaving us with 86 stores.
policies. Thus, they deemed the concept of having some employees on a different compensation plan to be extremely inappropriate.

Second, we needed to choose stores that showed similar characteristics to surrounding stores, but we also had to avoid choosing stores that were too close in geographical proximity to treatment stores so as to avoid the “water cooler effect.” That is, we did not want the salespeople in the control and treatment groups to communicate with each other, as such communication could potentially bias the outcome of the field experiment. For our control group, we ended up randomly selecting five stores (consisting of 26 salespeople) from a set of stores not in very close proximity to other stores, yet in the metropolitan areas of Stockholm, Gothenburg, and Malmö. To further avoid the water cooler effect, we made sure that there were no major sales training programs or conferences around the time of the field experiment, as these events could have potentially led to a spillover of information. Also, we made sure that there were no employee transfers between the treatment and control stores during the experiment. Indeed, the focal firm’s management was very concerned about the water cooler effect as well, but for a different reason. As mentioned above, the firm prided itself on its HR policy of one-for-all. This is the main reason that, for our control group, we were limited to only five stores for an observation period of one month.

---

15 Sweden is often cited as having one of the highest effective tax rates in the world. It is also known for having generous social security benefits, including child care, health care, housing allowances, and welfare. Sweden also ranks high with regard to gender equality (World Economic Forum, 2014). Hence, fairness is a relatively strong social norm in Sweden.

16 The term “water cooler effect” refers to the phenomenon in which employees gather around the office water cooler to talk. In our context, “water cooler effect” refers to the flow of information that potentially disrupts the motivation of the salespeople and, thus, the direction and effect size of the experiment treatments.
For our field experiment to be valid, it is important that the performance of the control and treatment groups is similar, especially related to the variability in productivity across time. We show in Table 2 the summary statistics of both groups in April 2015, the pre-intervention period. We see that both the mean and the standard deviation are quite similar across the two groups. Table 2 shows only that sales productivity for the control and the treatment groups are from similar distributions. However, and more importantly, for our empirical analysis to be valid, the temporal trend in the two groups must match. For verification, we perform several placebo tests in Section 6.

Table 3 shows the average SPH for the control and treatment groups across April and May 2015. Once again, the compensation plan changed from a monthly to a daily quota plan as of May 1. We can see the benefit and importance of having a control group. Given the 10% improvement in sales productivity (i.e., an increase in average SPH from $149.06 to $163.96) across the two periods for the treatment group, one would conclude that the plan change was immensely successful and that the daily quota plan outperforms the monthly quota plan. However, taking into account the differences in productivity of the control group across the two periods, the conclusion is not so obvious. There seems to be only a marginal effect—a 10.0% increase for the treatment group compared with a 9.1% increase (i.e., an increase in average SPH from $149.17 to $162.75) for the control group. The above analysis shows only an aggregate result of our field experiment, not taking into account individual heterogeneity (both in terms of permanent heterogeneity and responsiveness to incentives) and daily aggregate shifts in demand. We take
these features into account in our empirical analyses and model sales as a function of a salesperson’s effort, conditional on the compensation plan, in the following section.

5. Empirical Model

We model sales productivity $Y_{id}$ of salesperson $i$ on day $d$ as a multiplicative function of the salesperson-specific effects $\alpha_i$, common daily time trends $\gamma_d$, the compensation plan $z_{id}$, and an idiosyncratic shock $\varepsilon_{id}$ such that

$$Y_{id} = \exp(\alpha_i + \gamma_d + \delta z_{id} + \varepsilon_{id}),$$

(2)

where $z_{id}$ is a binary variable with a value of one if salesperson $i$ is in the treatment group and day $d$ is in the treatment period. The parameter $\alpha_i$ represents unobserved individual heterogeneity that is constant over time, and the parameter $\gamma_d$ represents any intertemporal variations that are common across all salespeople on a particular day. Examples include seasonal fluctuations in demand (e.g., due to the weather) or the firm’s other marketing activities (e.g., advertising) that affect all salespeople equally on day $d$. The parameter $\delta$ represents any increase (or decrease) in salesperson $i$’s effort as a result of the change in quota frequency of the compensation plan. The idiosyncratic shock $\varepsilon_{id}$ represents any other elements that affect sales, such as luck (either good or bad). We allow for heterogeneity in the idiosyncratic shocks and assume that these shocks are independently, identically, and normally distributed within salespeople over time with mean zero.
and variance $\sigma_i^2$. We take the logarithmic transformation of Equation (2) for our empirical model such that

$$\log(Y_{id}) = y_{id} = \alpha_i + \gamma_d + \delta z_{id} + \varepsilon_{id}.$$  (3)

With the presence of a control group, our identification of the treatment effect will result from any difference in productivity between the treatment group and the control group, allowing us to control for any natural trends common to both groups.\(^{17}\) Technically, we are able to identify the treatment effect just by cross-sectional analysis, using data only from periods after the treatment (May 1) if we assume homogeneity—that is, the sales of the treatment and control groups are identical before the treatment. Our sample size of 337 employees, although quite large for a field study, is not sufficiently large enough for random assignments in treatment conditions to wash away individual fixed effects. Our empirical approach allows the use of full information from the data to better control for individual heterogeneity, giving us more robust estimates of the treatment effect. We present the results of various model specifications in the next section.

6. Results

We first discuss the effect of quota frequency on overall sales performance. Next, we perform robustness tests to validate our empirical results. Then, we check for the effects of quota frequency on other dimensions of performance: the quality of effort (product returns) and the class of effort (product focus).

\(^{17}\) The difference-in-differences method (Card & Krueger, 1994) is frequently used to mimic an experimental research design with naturally occurring data.
6.1. The Effect of Quota Frequency on Sales Performance

The first column of Table 4 shows the result of Equation (3) with the logarithm of SPH per day as the dependent variable. Consistent with the model-free results in Table 3, in aggregate, we find no significant increase in productivity under the daily quota scheme compared to that of the monthly quota scheme. The second column of Table 4 shows the results of a deviant model of Equation (3). Specifically, to examine heterogeneity in responsiveness across agents, we allow for different slope parameters by segments of salespeople such that

$$\log(Y_{id}) = y_{id} = \alpha_i + \gamma_d + \sum_{r=1}^{R} \delta_r I_{(i \in S_r)} z_{id} + \epsilon_{id}, \quad (4)$$

where $I_{(i \in S_r)}$ is an indicator function that equals one if salesperson $i$ is a member of Segment $r$, $S_r$, and $\delta_r$ is the corresponding segment-level parameter. The segmentation is conducted via a quartile split with regard to sales productivity before the treatment period.\(^{18}\) We see an interesting pattern of responsiveness by salespeople of different productivity levels. While the quota frequency (daily quota plan) has a positive effect on less productive salespeople, it has a negative effect on more productive salespeople. In terms of magnitude, we find that the quota frequency has a positive and significant effect on the lowest-productivity quartile segment, with an 11.7% increase in sales productivity.\(^ {19}\) In contrast, we find a negative and significant effect of quota frequency on sales for

---

\(^{18}\) Because we use the difference in productivity from the control group to identify the treatment effect, we control for the regression-to-the-mean phenomenon in the data. This pattern in the data, if any, would be present for both the control and treatment groups. We thank an anonymous reviewer for this comment.

\(^{19}\) Because of the logarithmic transformation of our dependent variable, the magnitude of the estimated treatment effect is equal to an 11.7% increase in sales productivity, using the transformation formula $\exp(0.111)-1=0.1174$. 

26
the highest-productivity quartile segment—an 8.1% decrease in sales. Although directionally consistent with the above assessment, we don’t find any statistically significant effects for the two mid-quartile segments.

Frequent quotas (shorter-term goals) seem to have a positive effect on less productive salespeople. Why would this be the case? Less productive salespeople are, by definition, individuals who have greater disutility of effort or who are less efficient for a given amount of effort, or both. Hence, under a monthly quota plan, a salesperson who had bad luck (reduced sales) in the earlier part of the month will give up in the later days of the month because there is virtually no chance of meeting quota at the end of the month, given his or her low effort levels due to the high cost of effort. This would not be the case under a daily quota plan, as there is a fresh start every day, in which past performance does not affect current payoff and, thus, does not distort current motivation. This is consistent with our theoretical model in Section 3.

The education literature has found that frequent testing results in better outcomes (see Bangert-Drowns, Kulik, & Kulik (1991) for an extensive survey). Furthermore, studies in the psychology literature have found that breaking up a main goal into multiple subgoals results in more-favorable outcomes (Heath, Larrick, & Wu, 1999). Similarly, in the marketing literature, frequent goals have resulted in favorable outcomes (Gal & McShane; 2012, Zhang & Gao; 2016). In the sales management literature, Chung, Steenburgh & Sudhir (2014) explore the concept of quota frequency and find that the quarterly bonuses help salespeople to achieve the annual quota. However, their analysis was based on counterfactual simulations, using estimates from their
structural model, and not inferred directly from the data.\textsuperscript{20} To the best of our knowledge, sales quota frequency has never been directly analyzed from empirical data, let alone through a field experiment. Our results show that frequent quotas benefit less productive salespeople, much like the way in which frequent classroom testing helps improve the performance of lower-ability students.

In contrast to the findings with regard to less productive salespeople, frequent quotas lead to a decrease in performance for highly productive salespeople. To the best of our knowledge, this is a novel finding that has not been previously documented. It is, however, consistent with our theoretical prediction presented in Section 3. We discuss this result, in detail, in Section 6.4.

6.2. Robustness tests

As mentioned in Section 4, for our empirical analysis to be valid, the time trend must be parallel between the control and treatment groups. Because our control group was smaller than the treatment group and was chosen quasi-randomly (randomly chosen from a set of stores), we have to verify whether the parallel trend assumption is valid. For verification, we perform several placebo tests.\textsuperscript{21} These tests are conducted by pretending that a treatment occurred in a specific time for the treatment group during the non-treatment period (April). We conduct the placebo tests.

\textsuperscript{20} After estimating a structural model of heterogeneous sales force behavior responding to annual and quarterly quotas, Chung, Steenburgh, & Sudhir (2014) demonstrate, using counterfactual simulations, that removing quarterly quotas (and just keeping annual quotas) leads to a greater decrease in the performance of low-performing salespeople. In contrast, in this research, we directly test the effect of change in quota frequency on sales performance and provide direct empirical support for this conjecture. In addition, we investigate the effect of quota frequency not only on aggregate sales amounts, but also on various other dimensions of performance, such as the quality and type (class) of effort, by monitoring and measuring product returns and changes in the product focus.

\textsuperscript{21} We thank the Department Editor for suggesting this robustness check.
tests as follows. First, only using April data, we choose each week (1st, ..., 4th) and imagine that a treatment had been implemented for the treatment group. Then, we perform the same analysis using Equation (4). The results are shown in Table 5. No matter which week we use for the false treatment period, we see that all of the parameter estimates are very small and statistically insignificant. These results suggest that the treatment effect we find in our empirical analysis is a result of actual treatment and not that of any other systematic temporal changes, different from the treatment to the control groups.

6.3. The Effect of Quota Frequency on Product Returns

The focal firm prides itself on being known for its excellent customer service. It trains its salespeople to be knowledgeable in the entire range of its products’ technical specifications and their applications. Also, the firm continually ensures that salespeople regularly undergo a significant amount of customer service training well above the industry norm in the country. Hence, management was concerned that the daily quota plan, while potentially increasing short-term motivation, might be harmful if it resulted in salespeople aggressively selling unnecessary products to customers, which could, in turn, result in an increase in returned merchandise. This concern was further aggravated by the fact that the firm did not penalize salespeople for returns by reducing their compensation in either the daily or the monthly quota plans.

To examine whether there was a change in returns, we tracked all returned products and mapped them back to their original sales. We then computed the ratio of returns per sales amount to create the variable returns-to-sales (RTS) ratio. For example, if a salesperson sold $1,000 worth
of goods on April 1, $30 worth of which were eventually returned, the RTS ratio would be 0.03.
The RTS ratio, which is normalized by total sales, provides insights into problems that may be
associated with service quality. Table 6 shows the results of Equation (3) and Equation (4),
using the logarithm of the RTS ratio per day as the dependent variable. There seems to be no
meaningful effect as, in both the homogeneous and heterogeneous models, none of the parameters
is statistically significant. This suggests that an aggressive compensation plan alone does not
necessarily lead to a decrease in the quality of effort or to unethical behavior. It is likely a
combination of an aggressive compensation plan and other factors, such as a misguided company
culture, that drives such behavior (Zoltners, Lorimer, & Sinha, 2016).

6.4. The Effect of Quota Frequency on Product Focus

In Section 6.1, we showed that quota frequency (daily quota plan) can have a positive or a
negative effect on performance for low- and high-productivity salespeople, respectively. To better
understand the different behavioral changes by heterogeneous salespeople, we conduct the
following analyses. First, we run Equation (4) using the logarithm of the number of products
sold per hour as a dependent variable. The results are presented in the first column of Table 7.
We see that the number of products sold increased for the lowest-productivity segment under the
daily quota plan. This indicates that a part of the reason of why sales revenue increased under a
frequent quota plan is that the raw number of products sold increased.

What about the types of products sold? The second column of Table 7 presents the results of
Equation (4) with the logarithm of the average price of products sold per day as the dependent
variable. Here, we see a negative and significant effect for the highest-productivity segment and, although statistically insignificant, negative effects across all other segments. This indicates that salespeople (particularly high-performing ones) under a frequent quota plan—that is, given shorter split goals—tend to focus on the sale of low-ticket products. This result is consistent with the results of our theoretical model, presented in Section 3. The fact that salespeople are selling low-ticket products is not attractive to the firm, especially with regard to high-performing salespeople, who are supposed to focus on high-ticket, high-value-added products that bring more profits to the firm. To the best of our knowledge, there has been no research, to date, focused on the causal relation between frequency of goals and the type (product focus) and quality (product returns) of ensuing effort.

Overall, our results provide sound empirical evidence—supported by theory—of various effects of quota frequency that give significant substantive insights into the use of quotas. While having frequent quotas in the compensation plan may increase the absolute sales amount, especially for low-performing salespeople, frequent quotas tend to induce, particularly, the most productive salespeople to focus on low-ticket, low-margin items. This may result in a decrease in sales of high-value-added products, thus hurting the firm’s profits.

7. Conclusion

Monetary incentives, in the form of conditional payments based on performance, are one of the key instruments that organizations use to motivate their employees. These incentives are especially important in the domain of personal selling. A sales force compensation plan typically
consists of a fixed salary plus a variable payment conditional on the salesperson achieving a certain threshold of performance—a sales goal—referred to as a sales quota. Despite the common use of quotas, we do not fully understand their role, especially with regard to their temporal frequency. Hence, in this study, we examine the causal effect of quota frequency on various dimensions of performance for different types of salespeople.

We explore the role of quota frequency by first developing a theoretical model of salespeople’s behavior. Our model takes into account many realistic elements, such as salespeople’s multi-dimensional effort, heterogeneity in ability, product focus, and forward-looking behavior. To empirically validate our theoretical predictions, we collaborated with a major Swedish retail chain to run a field experiment, varying the compensation structure for full-time salespeople. Specifically, holding everything else constant, we changed the sales force compensation scheme from a monthly to a daily quota plan. Because the quota was in the form of average sales per hour, the only change was an increase in quota frequency. We also deployed a control group that consisted of salespeople who did not experience a change in compensation. We use the difference between the performance of salespeople who were subjected to the change (the treatment group) and that of salespeople who were not (the control group) to identify, as cleanly as possible, the effect of quota frequency on various dimensions of sales performance.

We find that an increase in quota frequency (the change from a monthly to a daily quota plan) increases sales performance, but mainly for low-performing salespeople, by preventing them from giving up when confronted with early negative sales shocks within a quota cycle. Under a daily quota plan, every day is a fresh start; thus, salespeople’s motivation is intact throughout the
month, whereas under a monthly plan, salespeople will give up in the later days of the month if they are too far away from—and, thus, have no chance of meeting—the quota set by the firm. In contrast, we find negative effects of quota frequency for high-performing salespeople, as they alter their product focus from high-ticket to low-ticket products, resulting in a decrease in sales of high-value-added products and, thus, hurting the firm’s profits. We do not find any evidence of an increase in product returns with the possibly more aggressive compensation plan (the daily quota plan), although returns were not penalized.

In summary, we use two methodologies to provide a comprehensive look into the role of quota frequency on various dimensions of salespeople’s performance. Our findings will be valuable for firms as they design their sales compensation plans. While reducing the time horizon for a quota setting can potentially motivate the less productive salespeople to higher performance, firms need to be careful with the unintended consequences of such a move for the high performers. More importantly, firms should also understand the overall impact of changing the time period for evaluating salespeople’s performance on the quality and type of effort such changes will induce.

There are some limitations to note. Because of concerns over fairness—that is, managers were concerned that some employees were treated differently from others—we were able to deploy a control group for only five weeks. Naturally, if we could maintain a control group that is similar in characteristics and restrict the flow of information about the treatment group for longer periods, we would be able to better analyze the long-term effect of quota frequency. Furthermore, our study was a one-time intervention, and, thus, we were not able to examine sequence or order effects. Finally, our venue was in Sweden, a country well known for its high tax rate and generous
social security programs. Accordingly, fairness and a sense of community there is a prominent social norm. One could speculate that our findings and the size of the effect would be more concrete in societies in which individualism is more the social norm. Although not addressed in this study, these topics would be exciting areas for future research.
Appendix. Proof of Proposition 1

We first show that $\mu_t \geq \mu_N$ for all $N$ with the additional assumption $p^H = 0$; this assumption will be dropped shortly. Under this assumption, 1) the agent never exerts effort on high-ticket products, and, hence, 2) $\mu_N$ is never larger than $p^I \chi^I$ for any $N \geq 1$, where the upper bound $p^I \chi^I$ is achieved if and only if the agent exerts effort (on low-ticket products) all the time.

Suppose that $N = 1$. If the agent exerts effort on low-ticket products, $X = \chi^L$, and, thus, the agent obtains the compensation $\beta(\chi^L)\chi^L$ with probability $p^I$. The compensation is zero with the complementary probability. Hence, it is optimal for the agent to exert effort on low-ticket products if and only if $p^I \beta(\chi^L)\chi^L \geq c^L$, and

$$
\mu_t = \begin{cases} 
    p^I \chi^L & \text{if } p^I \beta(\chi^L)\chi^L \geq c^L, \\
    0 & \text{if } p^I \beta(\chi^L)\chi^L < c^L.
\end{cases}
$$

Now, we show that $\mu_t \geq \mu_N$ for any $N > 1$. This is clearly true when $p^I \beta(\chi^L)\chi^L \geq c^L$ because in this case, $\mu_t$ achieves the upper bound $p^I \chi^I$. Next, consider the case $p^I \beta(\chi^L)\chi^L < c^L$. The agent exerts no effort when $N = 1$. Hence, it suffices to show that the agent also exerts no effort for any other $N > 1$. To show this, consider an alternative commission scheme $\tilde{\beta}$ such that $\tilde{\beta}(x) = \beta(\chi^L)$ for all $x \geq 0$. Note that $\tilde{\beta}(X / N) \geq \beta(X / N)$ with probability 1; hence, the agent exerts more effort only under the commission scheme $\tilde{\beta}$ compared to the case under $\beta$. However, if the agent exerts effort under the commission scheme $\tilde{\beta}$, the agent’s expected payoff would be

$$
p^I \tilde{\beta}(\chi^L)\chi^L - c^L = p^I \beta(\chi^L)\chi^L - c^L < 0.
$$
Hence, the agent never exerts effort under the commission scheme $\tilde{\beta}$ for any $N > 1$, so also does not under the commission scheme $\beta$.

Finally, the proof of the proposition is completed by observing that the above argument continues to be valid when $p''$ is positive but close to zero; the agent still never exerts effort on high-ticket products because doing so can never compensate disutility $c'' > 0$, and one can easily check that all the arguments above remain valid in this case.
References


Table 1: The Variable Compensation Plan

<table>
<thead>
<tr>
<th>Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quota ($sales/hour)</td>
<td>140</td>
<td>180</td>
<td>200</td>
<td>235</td>
<td>250</td>
</tr>
<tr>
<td>Commission rate (%)</td>
<td>0.27</td>
<td>0.67</td>
<td>0.9</td>
<td>1.5</td>
<td>2.0</td>
</tr>
</tbody>
</table>

The quota and commission rate numbers are approximate for confidentiality reasons.

Table 2: Summary Statistics by Group Pre-Intervention (April)

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>154.28</td>
<td>152.29</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>59.36</td>
<td>59.70</td>
</tr>
</tbody>
</table>

SPH per day across salespeople and time. Unit is in U.S. dollars.

Table 3: Average Sales per Hour (SPH) by Group across Periods

<table>
<thead>
<tr>
<th></th>
<th>April</th>
<th>May</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>149.17</td>
<td>162.75</td>
</tr>
<tr>
<td>Treatment</td>
<td>149.06</td>
<td>163.96</td>
</tr>
</tbody>
</table>

Average SPH is computed by summing up all sales and dividing them by total working hours. Unit is in U.S. dollars.

Table 4: The Effect of Quota Frequency on Sales Performance

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (Homogeneous)</th>
<th>Model 2 (Heterogeneous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily quota</td>
<td>-0.003 (0.031)</td>
<td>0.111 (0.039)</td>
</tr>
<tr>
<td>Daily quota—Q1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily quota—Q2</td>
<td></td>
<td>0.020 (0.038)</td>
</tr>
<tr>
<td>Daily quota—Q3</td>
<td></td>
<td>-0.038 (0.036)</td>
</tr>
<tr>
<td>Daily quota—Q4</td>
<td></td>
<td>-0.084 (0.036)</td>
</tr>
<tr>
<td>Agent fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Dependent variable: logarithm of sales-per-hour per day. Heteroscedasticity-consistent (Eicker–Huber–White) standard errors shown. Significance (at the 0.05 level) in bold.
### Table 5: Placebo Test

<table>
<thead>
<tr>
<th>Placebo week in April</th>
<th>week 1</th>
<th>week 2</th>
<th>week 3</th>
<th>week 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily quota—Q1</td>
<td>0.036</td>
<td>0.015</td>
<td>-0.055</td>
<td>-0.072</td>
</tr>
<tr>
<td></td>
<td>(0.083)</td>
<td>(0.070)</td>
<td>(0.073)</td>
<td>(0.069)</td>
</tr>
<tr>
<td>Daily quota—Q2</td>
<td>0.038</td>
<td>0.088</td>
<td>-0.055</td>
<td>-0.060</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.067)</td>
<td>(0.069)</td>
<td>(0.066)</td>
</tr>
<tr>
<td>Daily quota—Q3</td>
<td>0.020</td>
<td>0.039</td>
<td>-0.025</td>
<td>-0.041</td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td>(0.064)</td>
<td>(0.066)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Daily quota—Q4</td>
<td>-0.017</td>
<td>0.077</td>
<td>-0.002</td>
<td>-0.029</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td>(0.062)</td>
<td>(0.064)</td>
<td>(0.062)</td>
</tr>
</tbody>
</table>

| Agent fixed effects   | Yes    | Yes    | Yes    | Yes    |
| Time fixed effects    | Yes    | Yes    | Yes    | Yes    |

The Placebo tests are conducted using pre-intervention (April) data in which four separate weeks are used as imaginary treatment periods. Dependent variable: logarithm of sales-per-hour per day. Heteroscedasticity-consistent (Eicker–Huber–White) standard errors shown. Significance (at the 0.05 level) in bold.

### Table 6: The Effect of Quota Frequency and Returns

<table>
<thead>
<tr>
<th>Model 1 (Homogeneous)</th>
<th>Model 2 (Heterogeneous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily quota</td>
<td>-0.040 ( (0.110) )</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Dependent variable: logarithm of returns-to-sales ratio per day. Heteroscedasticity-consistent (Eicker–Huber–White) standard errors shown. Significance (at the 0.05 level) in bold.
Table 7: The Effect of Quota Frequency on Quantity and Price

<table>
<thead>
<tr>
<th>Quantile \ Dependent variable</th>
<th># products sold per hour</th>
<th>Average price of products sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily quota—Q1</td>
<td>0.128</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Daily quota—Q2</td>
<td>0.044</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Daily quota—Q3</td>
<td>-0.017</td>
<td>-0.023</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Daily quota—Q4</td>
<td>-0.041</td>
<td><strong>-0.047</strong></td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Agent fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Dependent variable: first column; logarithm of sales quantity per hour per day, second column; logarithm of average price of products sold per day. Heteroscedasticity-consistent (Eicker–Huber–White) standard errors used. Significance (at the 0.05 level) in bold.
Figure 1: Types of Variable Compensation Plans with Quotas

a) Commission at Quota

b) Bonus at Quota

c) Commission & Bonus at Quota
d) Commission & Bonus at Multi-tier Quota

Figure 2: The Misalignment of the Agent’s Effort with the Firm (N=1)

a) The Agent’s Policy

b) The Misalignment of Effort

Horizontally shaded area: agents who focus on high-ticket products.
Gridded area: agents who focus on low-ticket products.

Diagonally shaded area: agents who focus on low-ticket products but whose actions are suboptimal for the firm.
Gridded area: agents who focus on low-ticket products and whose actions are optimal for the firm.
Figure 3: The Relation between Quota Frequency and the Class of Effort

a) The Misalignment of Effort \((N=1)\)

\[ p^n = A(p^i) \]

Gridded area: agents who focus on low-ticket products but who are more profitable (to the firm) selling high-ticket products.

b) The Change in the Class of Effort \((N=2)\)

\[ p^n = A(p^i) \]

Dotted area: agents who focus on high-ticket products in the second period, conditional on positive sales in the first period.

Figure 4: Relation between Sales and Commission

a) Monthly Quota Plan (–April)

Illustrates monthly commission pay for a salesperson assigned 140 hours a month (April, 2015)

b) Daily Quota Plan (May–)

Illustrates daily commission pay for a salesperson assigned 7 hours a day (May, 2015)