Do Bonuses Enhance Sales Productivity?

A Dynamic Structural Analysis of Bonus-Based Compensation Plans

Doug J. Chung, Thomas Steenburgh and K. Sudhir*

September 2011

*Doug J. Chung is a PhD candidate in Marketing at Yale School of Management (doug.chung@yale.edu), Thomas Steenburgh is an Associate Professor of Marketing at Harvard Business School (tsteenburgh@hbs.edu), and K. Sudhir is James L. Frank Professor of Private Enterprise and Management (k.sudhir@yale.edu) at Yale School of Management.
This paper is the first essay of the first author's doctoral dissertation at Yale University.

Do Bonuses Enhance Sales Productivity?

A Dynamic Structural Analysis of Bonus-Based Compensation Plans

Abstract

We estimate a dynamic structural model of sales force response to a bonus based compensation plan. The paper has two main methodological innovations: First, we implement empirically the method proposed by Arcidiacono and Miller (2011) to accommodate unobserved latent class heterogeneity with a computationally light two-step estimator. Second, we estimate discount factors in a dynamic structural model using field data. The key to identification of discount factors is that bonuses affect only future payoff in non-bonus periods providing exclusion restrictions on current payoffs. Further, we exploit differences in predicted effort (and thus sales) over time from the exponential and hyperbolic discounting models to identify present bias in a hyperbolic discounting model. Substantively, the paper sheds insights on how different elements of the compensation plan enhance productivity. We find evidence that: (1) bonuses enhance productivity across all segments; (2) overachievement commissions help sustain the high productivity of the best performers even after attaining quotas; and (3) quarterly bonuses help improve performance of the weak performers by serving as pacers to keep the sales force on track to achieve their annual sales quotas. We also find clear evidence of hyperbolic discounting by salespeople.

Key Words: Sales force compensation, bonuses, quotas, dynamic structural models, two step estimation, discount factors, hyperbolic discounting.

1 Introduction

Personal selling is one of the most important elements of the marketing mix, especially in the context of B2B firms. An estimated 20 million people work as salespersons in the United States (Zoltners et al. 2008). Sales force costs average about 10% of sales revenues and as much as 40% of sales revenues for certain industries (Heide 1999). In the aggregate, U.S. firms spent over $800 billion on sales forces in 2006, a sum three times larger than advertising spending (Zoltners et al. 2008).

Marketing researchers routinely create response models for marketing mix instruments such as price, sales promotion and advertising. Meta-analysis of various research studies estimate that the sales force expenditure elasticity is about 0.34 (Albers et al., 2010), relative to about 0.22 for advertising (Assmus, Farley and Lehmann 1983) and about -2.62 for price (Bijmolt et al. 2005). While relative sales force expenditure elasticity is useful in determining the relative effectiveness of different instruments in the marketing mix, they give us little insight on how to design a sales force compensation plan, which is widely understood to be the primary tool by which firms can induce the sales force to exert the optimal levels of effort and thus to optimize the use of sales force expenditures.

A compensation plan can consist of many components: salary, commissions, and bonuses on achieving a certain threshold of performance called quotas. Figure 1 shows a variety of compensation plans that include combinations of these components. According to Joseph and Kalwani (1998), only about 24% of firms use a pure commission-based plan; the rest used some form of quotas. As per the Incentive Practices Research Study (2008) by ZS Associates, 73%, 85% and 89% in the pharma/biotech, medical devices and high tech industries respectively uses quota based compensation.

This paper has two substantive goals: First, to gain insight on how a firm should design its compensation plan. Should a firm offer quotas and bonuses in addition to commissions? Despite the ubiquity of quota-based compensation, there is considerable controversy in the theoretical (e.g., Holmstrom and Milgrom 1987; Lal and Srinivasan 1993) and empirical literature

(Oyer 1998; Steenburgh 2008) about whether quotas and bonuses are more effective than straight linear commission plans. Our paper sheds light on this issue by estimating a dynamic structural model of how the sales force responds to various compensation instruments such as commission rates, quotas and bonus levels.

Second, what should be the frequency of bonuses? Should one use a monthly, quarterly or annual bonus? Should one use a quarterly bonus in addition to an annual bonus? In the education literature, researchers have argued that frequent testing leads to better performance outcomes (Bangert-Drowns et al. 1991). Can quarterly quotas serve a similar role to improve outcomes? As in the education literature, where frequent exams help students to be prepared for the comprehensive final exam; frequent quota-bonus plans may serve as a mechanism to keep the sales force motivated to perform in the short-run well enough to be in striking distance of the overall annual performance quota.

Methodologically, the paper offers two key innovations. First, we incorporate unobserved heterogeneity in a latent class framework within a computationally light two step conditional choice probability (CCP) estimator for the dynamic structural model. Though the use of two step estimation approaches have recently gained popularity (Hotz and Miller 1993; Bajari, Benkard and Levin 2007), due to ease of computation relative to traditional nested fixed point estimation approaches (e.g., Rust 1987), their use in empirical applications has been limited by their inability to accommodate unobserved heterogeneity. Arcidiacono and Miller (2011) propose an approach that accommodates latent class heterogeneity within the two-step framework. To the best of our knowledge, ours is among the first empirical papers applying the Arcidiacono and Miller approach to account for unobserved heterogeneity in the two-step dynamic structural estimation framework.[1]

Second, and of importance to the dynamic structural modeling literature, we *estimate* rather than *assume* discount factors. It is well known in the literature on dynamic structural models

[1] Finger (2008) and Beauchamp (2010) are two concurrent working papers implementing this approach in economics.

that discount factors cannot be identified in standard applications because there are no instruments that provide exclusion restrictions across current and future period payoffs (Rust 1994). Hence the standard approach is to assume discount factors. Our bonus setting allows us to estimate discount factors. Since bonus payoffs occur only at the end of each quarter or year, in the non-bonus periods the probability of achieving quota and receiving bonus will not affect current payoff, but only future payoffs. Only a forward looking person (i.e., one with a non-zero discount factor) would respond to proximity to quota in non-bonus periods. We demonstrate through reduced form evidence that such behavior exists in the data, and then exploit this exclusion restriction to identify discount factors.

The constant exponential discounting (Samuelson 1937) model is the normative standard for inter-temporal behavior. Researchers in marketing and economics therefore routinely assume such behavior among agents in estimating dynamic structural models. Yet, the psychology and behavioral economics literature has shown strong evidence that hyperbolic discounting (Thaler 1981, Ainslie 1992, Laibson 1997) explains agent behavior better in many settings. Under hyperbolic discounting, agents discount the immediate future from the present more heavily than the same time interval starting at a future date. It is typically represented by the quasi-hyperbolic discount function $D(t) = \beta\delta^t$ (Phelps and Pollak (1968), Elster (1979), Laibson (1997, 1998) where $\beta < 1$ is the short-run *present bias factor* and δ is the *long-term discount factor*. When $\beta = 1$ (no present bias), it reduces to the exponential discount model. Hyperbolic discounting can lead to time inconsistent preferences and preference reversals. For example, a hyperbolic discounter can prefer 100$ today (t=0) to 120$ in a year (t=1), yet prefer 120$ in two years (t=2) to 100$ in one year (t=1).

For a given level of disutility for effort, the exponential and hyperbolic discount models differ in their predictions for inter-temporal effort in the bonus setting discussed above. Under exponential discounting, the agent's effort would be more smooth over bonus and non-bonus periods, while under hyperbolic discounting, the agent would concentrate effort in the bonus period (but less concentrated than a myopic agent). In our finite horizon setting, the utility function is identified off the last period, where there is no forward looking behavior and the problem reduces to a static model. Given knowledge of the utility parameters, we exploit the differences in predictions between hyperbolic and exponential discounting to estimate a present bias factor in addition to the long-term discount factor.

There are three specific modeling and estimation challenges in the structural estimation of response to compensation plans, especially those with quotas and bonuses. First, in a typical structural model, one observes the agent's action in response to the firm's action. For example in a consumer response model, one observes consumers' choices in response to the firm's choice of marketing mix such as price, advertising or sales promotion. In contrast, for a sales force response model, one does not observe the actions of the sales force, i.e., the exerted effort. One only observes the outcome of the agent's effort, i.e., sales, which is correlated with effort. Hence one has to make an inference about the agent's action (effort) that leads to sales from the observed realized sales. This requires a modeling assumption on the link between sales and effort.[2]

A second challenge is that unlike marketing mix variables that change over time, the compensation plan remains stationary over at least a year. With no variation in plans in the data, how can one estimate the responsiveness of the sales force to compensation? Here we draw on an empirical insight from Steenburgh (2008) that can help identify the sales force response, when the compensation plan involves payments for reaching quotas. In any given period, a sales agent's optimal effort depends on her state: how close the person is to achieving her quota. A sales agent may find it optimal to reduce effort when she is close or very far from achieving quota, but may stretch herself to reach the quota, when she has a moderate chance of achieving this quota. This implies that the optimal level of effort (and therefore sales) would vary from period to period as a function of the agent's state (distance to quota).

[2] The issue has parallels in empirical channel response models. For example, Sudhir (2001) makes an inference about manufacturer actions (wholesale prices) from the observed retail price and sales to infer competition between manufacturers.

A third issue follows from the discussion of the second. While quotas enable identification of sales force response, it also induces inter-temporal dynamics in optimal sales force response behavior. An agent has to be concerned not just with the current payoff when expending effort, but the future payoff that she can obtain by being in a more favorable state that can facilitate obtaining a bonus. This implies that the estimated structural model needs to account for forward-looking behavior on the part of sales agents. This requires a dynamic structural model

We estimate the dynamic structural model of sales force response to various features of the compensation plan using sales force output and compensation data from a Fortune 500 firm that sells office durable goods. This firm used Plan F in Figure 1. In addition, bonuses are provided at two different frequencies: quarterly and annual. As the compensation structure of the focal firm features almost all dimensions in typically used compensation plans, we observe how the sales force responds to these different dimensions of the plan. This rich plan provides us two key benefits: First, the presence of bonuses helps us to identify and estimate discount factors. Second, even though theoretically one can perform counterfactuals of any type of compensation plan if we can estimate structural parameters (other than discount factors) for a sales person with a less rich compensation plan, an analyst or manager should have greater faith in the counterfactuals, based on parameters that were estimated from observed responses to different elements of the compensation plan.

The rest of the paper is organized as follows. Section 2 discusses the related literature. Section 3 discusses the institutional details of the compensation plan at the firm, provides a numerical example to give intuition about how bonuses induce effort, and provides some model free evidence that facilitates model building. We present the model and the estimation methodology in sections 4 and 5. Section 6 discusses the estimation results and the counterfactual analysis. Section 7 concludes.

2 Related Literature

The literature review is in two parts: We begin with the discussion of the literature relating to the substantive issue of the role of nonlinearities such as quotas and bonuses in compensation plans. Following this, we discuss the empirical literature on structural estimation of worker productivity.

In the theoretical literature, Basu et al. (1985) apply the principal agent framework of Holmstrom (1979) and demonstrate that a combination of salary plus commission (usually nonlinear with respect to sales) will be optimal. Rao (1990) also shows a similar result on the optimality of nonlinear compensation plans. However, Holmstrom and Milgrom (1987) and Lal and Srinivasan (1993) question the need for nonlinear compensation schemes. Using the specific assumptions of linear exponential utility and normal errors (LEN) they show that a linear commission incentive scheme can achieve the best possible outcomes for the firm.

Yet, why do most firms have quota based compensation plans? Why are compensation plans nonlinear? Raju and Srinivasan (1996) suggest that even though a commission over quota plan may not be theoretically optimal, they provide the best compromise between efficiency and ease of implementation. Others argue that quota based plans offer high powered incentives that can motivate salespeople to work harder (e.g., Darmon 1997). Park (1995) and Kim (1997) demonstrate that a quota-bonus plan may lead to the first-best outcome, but in their framework, quota-bonus plan is just one of many possible plans that lead to first best outcomes. Oyer (2000) shows that when participation constraints are not binding, a quota-bonus plan with linear commissions beyond quotas can be uniquely optimal, because it can concentrate the compensation in the region of effort where the marginal revenue from effort minus the cost of compensation is maximized.

There is limited empirical work addressing this issue. Based on an analysis of aggregate sales across different industries in different quarters, Oyer (1998) concludes that the negative effects of quota based plans encouraging sales people to maneuver the timing of orders are greater than the benefits obtained from more effort. Steenburgh (2008) questions whether aggregate data can be used to reach this conclusion. Using individual sales performance data from the same firm used in this study (utilizes compensation plan F in Figure 1), he finds that the net improvement in revenues from effort dominates the inefficiencies induced by inter-temporal dynamic considerations.

Our work is related to several static structural models of worker behavior such as Ferrall and Shearer (1999) and Paarsch and Shearer (2000), who endogenize the optimal contract choice of the firm, given linear contracts. In contrast to these papers, we seek to understand the response to nonlinear incentives, which require us to model the dynamic response of sales agents. However, we do not model the contract choice, because we do not have data on selection across contracts.

Copeland and Monnet (2008) estimate a dynamic structural model of worker productivity in a check-sorting environment with nonlinear incentives; unlike sales force productivity, there is limited unobserved uncertainty in check sorting productivity. Much of the variation in productivity here can be explained by observed characteristics such as machine breakdowns etc.

A contemporaneous paper by Misra and Nair (2011) is closest to this paper in methods and substantive context. Both papers use the two-step estimation technique; however our paper innovates on two key dimensions. First, we accommodate latent class heterogeneity within the two-step estimation framework—an issue that has been an econometric challenge for the literature for close to two decades. Misra and Nair sidestep the unobserved heterogeneity issue by estimating each sales person's utility function separately.[3] Second, unlike Misra and Nair, who assume discount factors, we contribute to the broader dynamic structural modeling literature by estimating discount factors (specifically hyperbolic discount factors) using field data.

Substantively, in contrast to our focus on quotas with bonuses (plan F in Figure 1), Misra and Nair analyze quotas with floors and ceilings on commissions (plan D in Figure 1). They conclude that quotas reduce performance. This is because of two characteristics of their quotas:

[3] This is similar to the estimating individual level utility functions in conjoint analysis or scanner panel data, when there are a large number of observations per individual. Further, the approach requires that sales people will exert effort equally across all customers--an assumption they show is valid in their data, but unlikely to hold in general. Our latent class approach works in the more common situation where there are limited observations per individual.

First, the quota ceiling (beyond which sales people receive zero additional compensation) limits the effort of the most productive sales people who would normally have exceeded the ceiling. Second, the company followed an explicit policy of ratcheting quotas based on past productivity. This reduces incentives of sales people to work hard in any given period, because hard work is penalized through higher future quotas. In contrast, we find that quotas coupled with bonuses enhance performance. In the plan we consider, the company offers extra overachievement commissions for exceeding quotas and use a group quota updating procedure that minimizes ratcheting effects. Thus the two papers offer complementary perspectives that enhance our understanding about how quotas impact performance.

3 Institutional Details and Model-Free Evidence

We first describe the details of the bonus based compensation plan, followed by a numerical example to clarify how bonuses can help serve as a stretch goal, and induce inter-temporal effort. We then provide model-free evidence of forward looking behavior, seasonality etc.

3.1 The Compensation Plan

The focal firm under study is a highly regarded multinational Fortune 500 company that sells durable office products primarily using its own direct sales force. Each sales agent is given an "exclusive" territory; the firm traditionally does not encourage group work or team cooperation among the sales force. The firm also has an indirect sales force through "rep" firms who do not compete with the direct sales force. They are paid purely on commission, unlike the regular sales force. [4]

Our analysis focuses on sales performance data from 348 sales people from the regular sales force during the three year period 1999-2001. The firm's compensation structure follows the pattern in Plan F of Figure 1 and the details of the compensation schedule for the period of analysis are described in Table 1. For the indirect sales force, the firm only provided us with an

[4] Such rep firms are the focus of Jiang and Palmatier (2010).

index of revenues. The indirect sales force revenue (ISR) index is set to a base of 1 for January 1999. We provide descriptive statistics of the data in Table 2.

Every month, sales people receive a fixed monthly salary (average $3585) and a commission of 1.5% of revenues generated in that month. In the first three quarters, a quarterly lump-sum bonus of $1,500 is paid if the quarterly quotas are met. At the end of the year (i.e., end of the fourth quarter) an annual lump-sum bonus of $4,000 is paid if the annual quota is met. Further, an overachievement commission of 3% is paid for any excess revenues beyond the annual quota. There are no caps on revenues for which an agent could obtain commissions or overachievement commissions. Overall, for a salesperson that meets all quotas, the salary component will be roughly 50% of total compensation.

In building annual and quarterly quotas, the company uses internal metrics called "monthly allocated quotas" to its sales force (based on expected monthly revenues, given seasonality and territorial characteristics), though these are not used for compensation. We do not use these quotas for our modeling and estimation, but use them to benchmark performance in the reduced form analysis.

The most important element in performance evaluation within the firm is the annual quota; i.e., the firm views a salesperson as having a successful year if the annual quota is met. From Table 2, we see that sales people meet their annual quota roughly 50% of the time. Quarter 3 is the toughest quarter in that salespeople meet their third quarter quota only 42.8% of the time. This appears to be because third quarter targets tend to be more difficult than in other quarters (and even the annual quota) as indicated by the highest average of the ratio of quota to indirect sales index (88.2) for quarter 3.

3.2 Numerical Example

The numerical example is intended to illustrate: (1) that bonuses can be more efficient than pure commissions; (2) that a person's distance to quota can induce heterogeneity in effort; and (3) that hyperbolic and exponential discounters differ in inter-temporal effort (and generated sales).

Let the utility function of the salesperson that trades off effort (e) and income from sales (s), who has sold S units at the beginning of the new period be:

$$U(s,e,S) = -de^2 + rs + BI_{\{s+S \geq Q\}}$$

where $-d$ is the disutility parameter and r is the commission rate (d>0, r>0) and B is the bonus for reaching quota (Q). For illustration, assume a direct match between sales and effort, i.e. $s = e$.

We first illustrate the potential efficiency of bonuses with a static model. For simplicity assume $S = 0$. In the pure commission case with no bonus where $d = 1$, $r = 10$ and $B = 0$, the optimal effort is $e^* = 5$. In the bonus case, with $Q = 10$, and $B = 30$, the optimal effort is higher at $e^* = 10$ and the compensation cost to the firm is $130. To achieve the same level of sales and effort ($e^* = 10$) from a pure commission plan, the commission rate r has to increase to 20 and costs more for the firm at $200. Figure 2a illustrates these results graphically. Thus the quota-bonus plan is more efficient.

Second, we illustrate how bonuses can serve as a stretch goal for salespeople who are moderately far away from the quota, again with a static model. Let $d = 2$, $r = 10$, $Q = 10$, and $B = 30$. Consider three scenarios of distance to quotas: $S = 0$ (far away from quota), $S = 5$ (moderately close to quota) and $S = 7$ (close to quota). Figure 2b shows that the optimal effort levels are $e^* = 2.5$, 5 and 3 for the three cases. Thus the salesperson exerts maximum effort at $S =$ 5, when moderately far away from quota, all else being equal.

Third we illustrate how inter-temporal effort is affected by forward looking behavior--specifically differences in behavior between hyperbolic and exponential discounters in response to bonuses and a comparison with a myopic agent. To illustrate, assume there are three periods and the bonus is received at the end of the third period, based on cumulative performance. So the per-period utility function is as in equation (1) except that the bonus-term is included only in period three. Let d=1, r=10, B=100 and Q=30. Thus, if the individual reaches 30 in sales over the 3 periods, she receives a lump-sum bonus of 100. For the hyperbolic discounter, assume $\beta = 0.6$ and $\delta = 0.95$ over each period. To create the equivalent of discounting over the three periods, we assume $\delta = 0.7359$ for the exponential discounter. Figure 2c compares the discounting function and optimal effort levels for hyperbolic and exponential discounters. Due to the present bias, the hyperbolic discounter discounts much more the immediate future (period 1 to period 2) and therefore has a much sharper kink with lower levels of effort in the early periods (non-bonus periods) and substantially greater concentration of effort during the bonus period. The exponential discounter also puts maximum effort in the bonus period, but the effort is smoother and evenly spread across periods compared to the hyperbolic discounter. Consequently, the hyperbolic discounter accumulates lower sales in the non-bonus periods and needs to put in more effort to reach the quota in the last period.[5]

Overall, the total effort for forward looking agents is greater than for myopic agents, because myopic agents do not work towards the bonus in periods 1 and 2 and therefore end up in a poor state in period 3, impossible to make quota. Therefore in contrast to the hyperbolic and exponential discounters who reach quota with a total effort of 30; the myopic agent only responds to commissions and puts in an effort of 5 for all three periods for a total effort of 15.

3.3 Model Free Analysis

We consider three features of the data that informs model development. First, we look at the evidence of forward looking behavior induced by bonuses and hence the need to develop a dynamic model. Second, we consider seasonality in the data. Finally, we test for the possibility of sales substitution across quarters by sales agents.

Forward Looking Behavior

As discussed, bonuses provide an exclusion restriction in that it does not impact current payoffs, but only future payoffs. To the extent that a sales agent's sales performance is affected

[5] Such hyperbolic discounting is seen in many other circumstances where it is normatively optimal to put in smoothed level of effort over time, but nevertheless people concentrate their efforts around the deadline. For example, it is common for students to substantially expand study hours close to the examination, and not preparing enough earlier.

by variables relating to proximity to bonuses, this is evidence of forward looking behavior.[6] But proximity to bonus quota will impact performance only if agents have a reasonable chance of making quota. Figure 3 shows the graph of probability of reaching annual quota, conditional on the cumulative fraction of annual quota ($\%AQ$) achieved till November. It is clear that there is very little chance of achieving quota if $\%AQ < 0.5$. We therefore divide agents by their state $\%AQ < 0.5$ and %AQ > 0.5 to test if the state affect sales and estimate regressions on sales performance in November as a function of their state. Table 3 reports the results of the regressions. Consistent with forward looking behavior, the state $\%AQ$ is significant only for agents with $\%AQ > 0.5$.[7]

For additional evidence of forward looking behavior, we show scatter plots and the best fitting nonparametric smoothed polynomial (and its 95% confidence interval) of sales revenues normalized by monthly allocated quotas in the quarterly bonus months (March, June, September, December) against percentage of quota attained by the previous month in Figure 4a. For March, June and September, the x axis is the percentage of quarterly quota completed ($\%QQ$), while for December, the x axis is the percentage of annual quota completed ($\%AQ$). The vertical dotted line shows the $\%QQ$ and $\%AQ$ at which the salespeople on average achieve their monthly allocated quotas.

Two key elements stand out from Figure 4a. First, across the board there is little reduction in effort when salespeople are close to achieving quota, due to the overachievement commission rate. Second, there is a steady increase over time in the $\%QQ$ and $\%AQ$ threshold beyond which sales people reach their monthly targets. The threshold is about 25% in March, 35% in June, 45% in September and close to 70% in December. Early in the year, even if below targets, salespeople still have hopes of receiving the large annual bonus by working hard and with some good luck. As it gets closer to year-end, chances of reaching quota becomes less likely, and sales people respond by reducing effort even at higher levels of $\%AQ$ and $\%QQ$. As annual bonuses

[6] A similar argument is made in providing evidence of forward looking behavior by students in the textbook market by Chevalier and Goolsbee (2009).

[7] The qualitative conclusions are robust in the range of thresholds of $\%AQ$ from 0.4 to 0.6. Note that if there were no overachievement commission, agents very close to quota may reduce their effort.

should have no impact on current payoffs in March, June or September, but only on future payoffs, this is suggestive of forward looking behavior.[8]

The next set of graphs presented in Figure 4b, shows the same relationship in the pre-bonus months (February, May, August and November) and provides additional evidence for forward looking behavior. In the early months, February, May and even August, at all levels of $\%QQ$, the salesperson on average sells above the monthly allocated quota. This is because hard work (and some good luck in the form of positive sales shocks) may give a reasonable chance of attaining the smaller quarterly targets. However, in November, only at a very high level of $\%AQ$, does the salesperson sell above the monthly allocated quota, because one has a very limited chance of making up the large gap in just two months. In the pre-quarterly bonus months, the immediate future quarterly bonus impacts behavior, even though it has no impact on current payoff; indicating more conclusive forward looking behavior.

This evidence leads to a natural question. Should the large annual bonus be split into a quarterly bonus (as in other months) and an annual bonus? The quarterly bonus can prevent sales people from giving up in November, even if they do not have a chance of reaching the annual quota. But with such a quarterly quota, early in the year, agents may have limited incentive to stretch after reaching quarterly quotas. How these two issues tradeoff is an empirical question, which we subsequently address in the counterfactual analysis.

Seasonality

Figures 5a graphs the average revenues over the months for the regular sales force. There are clear peaks at the end of each quarter. These peaks could be either due to seasonality or bonuses at the end of each quarter. Figure 5b shows the index of indirect sales revenue (ISR) for the pure

[8] An alternative explanation is that targets in the early quarters are easier to achieve than those in later quarters. From the ratio of quarterly quotas to indirect sales index for the two quarters in Table 2, it is indeed true that September and December have higher ratios. Thus the two explanations are confounded in September and December. However quotas in the second quarter are less onerous than in the first quarter; yet agents give up more in the second quarter, this is consistent with our explanation of forward looking.

commission indirect sales force. As the ISR index is not contaminated by bonuses, we use it to control for seasonality and isolate the effect of bonuses on sales person effort and revenues.

To build intuition for how ISR can help control for seasonality and isolate effort, see Figure 5c which graphs direct sales force average revenues and multiples of ISR. At a multiple of around 50, the ISR virtually mimics the average revenues, making the revenues from the commissioned and bonus sales force close to identical. This suggests that bonuses are not effective in inducing additional effort. When ISR has a multiple of 30 or 40, even after the overall seasonality is accounted for, there is gap in revenues that we interpret as induced by effort. It is interesting that these gaps are larger at the end of the quarter, suggesting the value of bonuses in inducing effort. We empirically estimate the multiple for ISR in order to control for seasonality.[9]

Sales substitution across months

One possibility is that sales people giving up at the end of the quarter may be doing so to increase the odds of hitting quotas in subsequent quarters by simply not booking the sales in the current quarter. If this were true, then one should see a negative linkage between sales in months t and $t+1$; and especially between the last month of a quarter and first month of the next quarter. See Table 4 for regression results of revenue in month t against revenues in month $t-1$. The coefficient of first month of each quarter captures potential borrowing effects from the last month of a bonus period to the first period of the next bonus period. The first month of quarter is not significant suggesting little substitution across quarters (Models 1 and 2). It is not significant even if we separate the effect for people who are "way off target" in the last month and have therefore the greatest incentives to postpone purchases, as seen in Model 3.[10] We therefore do not model substitution across quarters.

[9] From Table 6a, we know the ISR multiple for our model is about 30, given the average of lagged annual quota is 1639.

[10] We defined "way off target" as those whose previous quarter sales were less than 50% of their quota. The results were robust and did not vary with alternative definitions of "way off target"..

4 Model

Based on the model-free evidence, we build a dynamic model of sales force response to the quota-based compensation scheme. The timing of the model is as follows:

1. At the beginning of each year, firm chooses the annual compensation plan.
2. Each month, agents observe their current state and exert effort in a dynamically optimal manner.
3. An idiosyncratic sales shock is realized; the shock plus agent's effort determines the agent's realized sales for the period. Agent receives compensation.
4. The realized sales of the current period affect the agent's state of the next period. Steps 2-3 are repeated each month until the end of the year and steps 1-3 are repeated over the years.

We describe the model in five parts: (i) the compensation plan (ii) the sales agent's utility function (iii) the state transitions (iv) effort as a function of state variables and (v) the optimal effort choice by the sales agent.

4.1 Sales Response Model

We model the sales revenue function (S_{it}) for salesperson i at time t in two parts: (1) a base level of sales independent of effort, parameterized by demand shifters (z_{it}^D) and (2) sales induced due to effort (e_{it}) parameterized by effort shifters that include territory and salesperson characteristics (z_{it}^E).

$$S_{it} = f(z_{it}^D) + e_{it}(z_{it}^E) + \varepsilon_{it} \tag{1}$$

where ε_{it} is an additive sales revenue shock, not anticipated by the salesperson when choosing effort.

As discussed, the market potential varies across territories and across time. To account for the cross-sectional variation in market potential, we use annual quota from the previous year (AQ_{iy-1}). To account for seasonality of demand across months, we use the ISR index, (ISR_t).

We also include an interaction between the two variables to allow for seasonality to have a larger impact on larger territories.

For effort shifters in e_{it}, we use the following variables: Given that effort is a function of demand shifters, we include both AQ_{y-1} and ISR_t in z_{it}^E. As discussed in the motivation, the salesperson's state with respect to achieving quota will have an impact on the effort they expend. We therefore use the cumulative percentage of quarterly and annual quota completed till time t ($\%QQ_{it}$, $\%AQ_{it}$) as variables that affect effort. In addition, we allow a time-invariant salesperson specific variable, tenure with the firm (τ) to moderate the level of effort. We allow for interactions between ISR_t and $\%QQ_{it}$, $\%AQ_{it}$ in the effort function, thus effort policy function is different for different months.[11]

Note that unlike the demand shifter function f, which is common across all salespeople, the effort function will vary across salespeople. Specifically, we allow for salespeople to belong to one of multiple discrete segments, hence these effort functions will be estimated at the segment level. We estimate the effort function non-parametrically, by using Chebyshev polynomials of the variables described above.

4.2 Compensation Plan

The compensation plan has three components. They are: (1) the monthly salary w_{it}, (ii) end-of quarter bonus, B_{iqt} for achieving the corresponding quarterly quota Q_{iqt}, and end of year bonus B_{iyt} for achieving the corresponding annual quota Q_{iyt} (3) commission rate r_{it} per dollar worth of sales and an overachievement commission rate, $r_{it}^{'}$ given at the end of the year for sales over and above the annual quota for each individual i at time t. We represent the compensation plan for a salesperson i by the vector $\psi_{it} = \{w_{it}, Q_{iqt}, Q_{iyt}, B_{iqt}, B_{iyt}, r_{it}, r_{it}^{'}\}$.

4.3 Sales person's per-period utility

[11] While it would be ideal to estimate separate effort policy functions for each month, there are not typically enough degrees of freedom in the data to do this for each month. We balance the flexibility/degrees of freedom tradeoff with our approach.

In each period t, sales person i receives positive utility of wealth W_{it} earned based on realized sales and a disutility $C(e_{it};\theta_i)$ from exerting effort e_{it}. Thus the utility function is defined as:

$$U\left(e_{it}, S_{it}; \psi_i, \theta_i, \gamma_i\right) = E\left[W\left(S_{it};\ \psi_i\right)\right] - \gamma_i \operatorname{var}\left[W\left(S_{it};\ \psi_i\right)\right] - C\left(e_{it}; \theta_i\right)$$ [12]

where γ_i and θ_i are the risk aversion and disutility parameters respectively for salesperson i.

Given the sales levels, and the compensation plan, the wealth for individual i, W_{it} can be computed. W_{it} arises from four components, the per-period salary component w_{it}, the lump-sum bonus component B_{it}, the commission component C_{it}, and the overachievement commission component OC_{it}. The detailed expressions of wealth is as follows,

$$W_{it} = w_{it} + B_{it} + C_{it} + OC_{it}$$

$$B_{it} = I_{qt} I\left(z_{i1t} + \frac{s_{it}(e_{it}(z_{it}^E), z_{it}^D; \alpha_i) + \varepsilon_{it}}{Q_{iqt}} > 1 \right) B_{qt} + I_{yt} I\left(z_{i2t} + \frac{s_{it}(e_{it}(z_{it}^E), z_{it}^D; \alpha_i) + \varepsilon_{it}}{Q_{iyt}} > 1 \right) B_{yt}$$

$$C_{it} = (s_{it}(e_{it}(z_{it}^E), z_{it}^D; \alpha_i) + \varepsilon_{it}) r_{it}$$

$$OC_{it} = I_{yt} I\left(z_{i2t} + \frac{s_{it}(e_{it}(z_{it}^E), z_{it}^D; \alpha_i) + \varepsilon_{it}}{Q_{iyt}} > 1 \right) \left(z_{i2t} Q_{iyt} + s_{it}(e_{it}(z_{it}^E), z_{it}^D; \alpha_i) + \varepsilon_{it} - Q_{iyt} \right) r_{it}^{'}$$

where z_{i1t} and z_{i2t} are the percentage of quarterly and annual quotas completed respectively by salesperson i until time t. I_{qt} and I_{yt} are indicators for whether time t is a quarterly or annual bonus period.

In our empirical analysis, we use a quadratic functional form for the disutility function; specifically, $C\left(e; \theta_i\right) = \theta_i e^2$. Thus the set of structural parameters of the salesperson's utility function that needs to be estimated are $\mu_i = (\theta_i, \gamma_i)$.

4.4 State Variables

[12] In the case of the CARA utility function (exponential utility function) with normal errors and a linear compensation plan, this functional form represents the certainty equivalent utility of the agent. Here we consider the utility function to be a second order approximation to a general concave utility function.

As discussed, the nonlinearity of the compensation scheme with quotas and bonuses introduces dynamics into the sales agent's behavior because there is an additional tradeoff between the disutility of effort today and a higher probability of lump-sum bonus and overachievement commissions tomorrow. To incorporate the dynamics of the model we consider the following stochastic state variables, the percentage of annual quota completed, the percentage of quarterly quota completed. These state variables evolve as follows:

1. Percentage of quarterly quota completed ($\%QQ$)

$$z_{i1t} = \begin{cases} 0, & \text{if } t \text{ is start of quarterly quota period} \\ z_{i1(t-1)} + \dfrac{S_{i(t-1)}}{Q_{iqt}}, & \text{otherwise} \end{cases}$$

2. Percentage of annual quota completed ($\%AQ$)

$$z_{i2t} = \begin{cases} 0, & \text{if } t \text{ is start of annual quota period} \\ z_{i2(t-1)} + \dfrac{S_{i(t-1)}}{Q_{iyt}}, & \text{otherwise} \end{cases}$$

Other state variables would include time varying demand shifters, ISR index and territory characteristics, for which we use previous year's annual quota. Naturally, the time varying indirect sales is a one-to-one mapping to period type and hence includes information about different periods. We use tenure with the focal firm (τ) as an individual state variable that impacts effort. These state variables are collected in a state vector $z_{it}^E = \left\{ z_{i1t}, z_{i2t}, IS_t, AQ_{i(y-1)}, \tau_i \right\}$.

4.5 Optimal Choice of Effort

Given the parameters of the compensation scheme ψ, and the state variables and their transitions, each sales agent would choose an effort level conditional on her states to maximize the discounted stream of expected future utility flow. Alternatively, if this value function is below the reservation wage, the salesperson may choose to leave the firm.

The stream of utility flow, under the optimal effort policy function, conditional on staying at the firm, and the behavioral notion of quasi-hyperbolic discounting can be represented by a value function,

$$V\left(z;\psi,\Omega\right)=\max_{e}U\left(e,z;\psi,\mu\right)+\beta\delta E\left[\max_{e'}U\left(e',z';\psi,\mu\right)+\delta E\left[\max_{e''}U\left(e'',z'';\psi,\mu\right)+\ldots\right]\right]$$

where $\Omega=\left(\mu,\beta,\delta\right)$ is the set of primitives or structural parameters of the underlying utility function μ, and the discount parameters β and δ. The long-term discount factor is δ, the short-term discount factor is $\beta\delta$, where $\beta<1$ indicates present bias. If $\beta=1$, the model reduces to a single parameter exponential discount model. The expectation of the value function is taken with respect to both the present and future sales shocks.

5 Estimation

Traditionally, the nested fixed-point algorithm (NFXP) developed by Rust (1987) is used to estimate dynamic models. However, NFXP estimators are computationally burdensome as one has to solve the dynamic program numerically over each guess of the parameter space for every iteration. The two-step estimation first introduced by Hotz and Miller (1993) and extended by Bajari, Benkard, and Levin (2007) can serve to reduce the computational burden. In this approach, the model estimation proceeds in two steps. In the first step, we estimate the conditional choice probabilities of choosing a certain action as a flexible nonparametric function of state variables. Then, in the second step, these conditional choice probabilities are used to estimate the structural parameters of the sales agent's utility function.

Until recently, it was believed that the accurate estimation of conditional choice probabilities for an agent is impractical when there is unobserved heterogeneity. Arcidiacono and Miller (2011) propose an EM–Algorithm based approach to accommodate unobserved heterogeneity in the first step of the two step estimation procedure. We provide one of the first applications of this approach – illustrating the empirical validity of the approach in practical applications. We now discuss the details of the two step estimation procedure.

5.1 Step 1

In this step, we need to estimate a flexible non-parametric mapping between observable states and actions of the sales person; this requires a non-parametric model of the monthly effort function $e_i(z_{it}^E)$, that links effort and state in equation (1). We model the effort function non-parametrically as a combination of basis functions of the state variables. Thus the non-parametric effort function is:

$$e_{it} = \sum_{\ell=1}^{L} \rho_\ell(z_{it}^E)\lambda_{i\ell} \tag{2}$$

where the ℓ^{th} basis function is $\rho_\ell(z_{it}^E)$. In this application, the ℓ^{th} basis function is the ℓ^{th} order Chebyshev polynomial.

From equations (1) and (2) we have the following sales response function to estimate.

$$S_{it} = f(z_{it}^D) + \sum_{\ell=1}^{L} \rho_\ell(z_{it}^E)\lambda_{i\ell} + \varepsilon_{it}$$

For z_{it}^D which is a subset of z_{it}^E, (from now on referred to as z_{it}), we use two variables: (1) lagged annual quota for salesperson i, (2) the indirect sales force revenue (ISR) index. We use the direct linear effect of these variables to control for cross sectional variations of territory characteristics and temporal variations in monthly seasonality. The interaction effects of these variables with the other state variables go into the polynomial function in (2).

The lagged annual quota takes into account general territory characteristics that are likely to be generated with limited effort, i.e., market size. The revenues from the indirect sales force capture market seasonality, independent of the nonlinear nature of the compensation plan. We assume that the revenue shocks (ε_{it}), come from an i.i.d. normal distribution.

If one could estimate the sales response and effort response function at the level of each individual, we can simply obtain the individual level parameters of the effort and sales policy function by maximizing the log likelihood of the sample such as

$$\hat{\Theta}_i = \arg\max \sum_{t=1}^{T} \log \left\{ L_i \left(S_{it} - f(z_{it}^D;\alpha_i) - \sum_{\ell=1}^{L} \rho_\ell \left(z_{it}^E \right) \lambda_{i\ell} \right) \right\} \tag{3}$$

where the vector $\hat{\Theta}_i = \left\{ \alpha_i, \lambda_i, \sigma_i \right\}$ contains the set of parameters of the sales response and effort policy functions and the distribution of sales shocks, where

$$L_i\left(\varepsilon\right) = \frac{1}{\sigma_i \left(2\pi\right)^{1/2}} e^{-\frac{1}{2}\left(\frac{\varepsilon}{\sigma_i}\right)^2} \tag{4}$$

We accommodate unobserved heterogeneity by allowing for discrete segments. Assume that sales person i belongs to one of K segments, $k \in \left\{1,\ldots,K\right\}$ with segment probabilities $q_i = \left\{ q_{i1},\ldots,\ q_{iK} \right\}$. Let the population probability of being in segment k be π_k. Let $\mathcal{L}(S_{it} \mid z_{it}, k; \Theta_k)$ be the likelihood of individual i's sales being S_{it} at time t, conditional on the observables z_{it}, and the unobservable segment k, given segment parameters Θ_k. Then the likelihood of observing sales history S_i over the time period $t = 1\ldots T$, given the observable history z_i, and the unobservable segment k is given by:

$$L_k\left(S_i \mid z_i; \Theta_k, \pi_k \right) = \pi_k \left(\prod_{t=1}^{T} \mathcal{L}_{ikt} \right) \tag{5}$$

where $\mathcal{L}_{ikt} = \mathcal{L}(S_{it} \mid z_{it}, k; \Theta_k)$. As noted earlier we assume the distribution of the revenue shocks to be normally distributed and hence use the normal likelihood for equation (5) as in equation (4). The parameter $\Theta_k = \{\alpha_k, \lambda_k, \sigma_k\}$ is the vector of segment level parameters of the sales response and effort policy function where each λ_k is the parameters that index the effort policy for segment k and σ_k is parameter for the distribution of the revenue shocks for segment k.

By summing over all of the unobserved states $k \in \left\{1,\ldots,K\right\}$, we obtain the overall likelihood of individual i:

$$L\left(S_i \mid z_i;\Theta,\pi\right)=\sum_{k=1}^{K}L_k\left(S_i \mid z_i;\Theta_k,\pi_k\right)$$

and hence the log-likelihood over the N sample of individuals becomes

$$\sum_{i=1}^{N}\log\left(L\left(S_i \mid z_i;\Theta,\pi\right)\right)=\sum_{i=1}^{N}\log\left(\sum_{k=1}^{K}\pi_k\prod_{t=1}^{T}\mathcal{L}_{ikt}\right) \tag{6}$$

Directly maximizing the log-likelihood in (6) is computationally infeasible because the function is not additively separable so we use the approach of Arcidiacono and Jones (2003) and Arcidiacono and Miller (2010) to iteratively maximize the expected log-likelihood in equation (7)

$$\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{t=1}^{T}q_{ik}\log\mathcal{L}\left(S_{it} \mid z_{it}\,k;\Theta_k\right) \tag{7}$$

where q_{ik} is formally defined below as the probability that individual i is of segment type k given parameter values $\Theta=\{\Theta_1,\ldots,\Theta_K\}$, where $\Theta_k=\{\alpha_k,\lambda_k,\sigma_k\}$ and segment probabilities $\pi=\{\pi_1,\ldots,\pi_k\}$, conditional on all of the observed data of individual i.

$$\Pr\left(k \mid S_i,z_i;\Theta,\pi\right)=q_{ik}\left(S_i,z_i;\Theta,\pi\right)=\frac{L_k\left(S_i \mid z_i;\Theta_k,\pi_k\right)}{L\left(S_i \mid z_i;\Theta,\pi\right)} \tag{8}$$

The iterative process is as follows: We start with an initial guess of the parameters Θ^0 and π^0. A natural candidate for such starting values would be to obtain the parameters from a model without unobserved heterogeneity and slightly perturbing those values.[13] Given the parameters $\left\{\Theta^m,\pi^m\right\}$ from the m^{th} iteration, the update of the $(m+1)^{\text{th}}$ iteration is as follows

a) Compute $q_{ik}^{(m+1)}$ using equation (8) with Θ^m and π^m

b) Obtain $\Theta^{(m+1)}$ by maximizing (7) evaluated at $q_{ik}^{(m+1)}$

c) Update $\pi^{(m+1)}$ by taking the average over the sample such that

[13] We started the initial values from one tenth of the standard error from the parameter values obtain from a single segment model. The initial values of the segment probabilities were set equally across segments.

$$\pi_k^{(m+1)} = \frac{1}{N}\sum_{i=1}^{N} q_{ik}^{(m+1)}$$

We iterate (a) through (c) till convergence.

For the basis functions in the effort policy, we use Chebyshev polynomials of state variables to approximate effort.[14] From the estimation, we obtain the vector of parameters for the basis functions (λ), the vector of parameters for the sales policy (α), and the parameters of the revenue shocks (σ) for each segment k. Also we obtain the population segment probabilities (π) for each segment. The procedure gives us the sales revenue function $\hat{S}(.)$ and effort policy function $\hat{e}(.)$ for each segment.

5.2 Step 2

The key idea of the two-step estimation is that in the first stage we observe the agent's optimal actions. Using these observed optimal actions we are able to construct estimates of the value function, which enables us to estimate the primitives of the model that rationalize these optimal actions.

Let the value function of a representative agent at state z that follows an action profile e, conditional on the compensation plan ψ, the sales profile S and the primitives of the utility function and discount parameters $\Omega = \left(\mu, \beta, \delta\right)$ be represented as

$$V\left(z; e; \psi, S, \Omega\right) = E\left\{\sum_{t=0}^{T} D(t) U(e(z_t), z_t, \varepsilon_t; \mu) \mid z_0 = z; \psi, S, \Omega\right\} \tag{9}$$

where $D(t) = \begin{cases} 1, & if\, t = 0 \\ \beta\delta^{\mathrm{t}}, & \text{otherwise} \end{cases}$ is the hyperbolic discount function, and the expectation operator would be over the present and future sales shock ε_t.

[14] For reference, see "Numerical Methods in Economics", Kenneth L. Judd, MIT Press, 1998.

Using the estimated sales and effort policy function and the distribution of the sales shocks in the first stage, we are able to forward simulate the actions of sales agents to obtain the estimate of the value function. The detailed simulation procedure is as follows.

a) From initial state of z_t calculate the optimal actions as $e(z_t)$

b) Draw sales shock ε_t from $f(\varepsilon)$

c) Update state z_{t+1} using the realized sales $s(e(z_t))+\varepsilon_t$

d) Repeat (a)-(c) until $t=T$

By averaging the sum of the discounted stream of utility flow over multiple simulated paths we can get the estimate of the value function $\tilde{V}\left(z;e(z);\psi,S,\Omega\right)$.[15]

Let $e^s(z)$ be any deviation policy from a set of feasible policies that is not identical to the optimal policy and, by using the same simulation method proposed above, let the corresponding estimate of the value function be called the sub-optimal value function $\tilde{V}\left(z;e^s\left(z\right);\psi,S,\Omega\right)$. Since $e(z)$ by definition is the effort policy and thus at an optimum, then any deviations from this policy rule would generate value functions of less or equal value to that of the optimal level.

Let us define the difference in the two value functions as,

$$Q(v;\psi,S,\Omega) = V(z;e(z);\psi,S,\Omega) - V(z;e^s(z);\psi,S,\Omega)$$

where $v \in \mathcal{V}$ denotes a particular $\{z,\ e^s(z)\}$ combination.[16] Then if $e(z)$ is the optimal policy, the function $Q(v;\psi,S,\Omega)$ would always have value of greater or equal to zero. Thus our estimate of the underlying structural parameters Ω would satisfy,

$$\hat{\Omega} = \arg\min \int (\min\left\{Q(v;\psi,S,\Omega),0\right\})^2 dH(v)$$

[15] For each segment, we drew four hundred simulation draws over each period and computed the value functions.

[16] As indicated in Bajari, Benkard, and Levin (2007), there are multiple ways to draw these suboptimal policy rules. Although the method of selecting a particular perturbation will have implications for efficiency the only requirement necessary for consistency is that the distribution of these perturbations has sufficient support to yield identification. We chose to draw a deviation policy from a normal distribution with mean zero and quarter of the variance from the revenue shock distribution, i.e. $e^s(z)=e(z)+\eta$.

where $H(v)$ is the distribution over the set $\mathcal{V}$ of inequalities. Our empirical counterpart to $Q\left(v;\psi,S,\Omega\right)$ would be $\tilde{Q}\left(v;\psi,\hat{S},\Omega\right)=\tilde{V}(z;\hat{e}(z);\psi,\hat{S},\Omega)-\tilde{V}(z;\hat{e}^{s}(z);\psi,\hat{S},\Omega)$ and as a result our estimates of the structural parameters are obtained from minimizing the objective function in equation (10).[17]

$$\frac{1}{N_I}\sum_{j=1}^{N_I}\left(\min\left\{\tilde{Q}(v_j;\psi,\hat{S},\Omega),0\right\}\right)^2 \tag{10}$$

The above procedure is performed for each segment with the segment specific effort policies obtained in Step1. This allows us to estimate the structural parameters for each segment.[18]

5.3 Identification

There are three major identification challenges. First, we do not observe effort. Hence the link between effort and sales cannot be identified non-parametrically. Second, in dynamic structural models, it is typically impossible to identify discount factors separately from the utility function. Third, we estimate hyperbolic discount factors. Below, we discuss how we address these issues.

Realized sales are a function of demand shifters, effort and additive sales shocks. Conditional on observed demand shifters and given multiple observations of sales at different states, we can separately identify non-parametrically the density of sales shocks and a deterministic function of effort. We assume a deterministic (but flexible) relationship between effort and observable states (%QQ and %AQ and demand shifters) for each segment. Finally, as we do not observe effort, we need a strictly monotonic parametric relationship between sales and effort. As we estimate a flexible relationship between observable states and effort, we model the relationship between sales and effort to be linear.

The discount factor is not identified separately from the utility function in standard dynamic structural models because typically there are no variables that do not affect contemporaneous

[17] We drew two hundred deviation strategies to construct the objective function and hence N_I=200.

[18] In addition, we used a second set of moment inequalities to reflect the participation constraint that employees continued to work at a firm because they at least obtained a reservation value (normalized to zero); i.e., $\min\{V(z;e(z);\psi,S,\Omega),0\}$. It turns out these inequalities are non-binding and do not impact our estimates.

utility, but only future utility. (Rust 1994; Magnac and Thesmar 2002). In the absence of such an exclusion restriction, this implies that if an agent exerts low effort in a period, it is not possible to distinguish if this is due to high disutility for effort, or because they discount future utilities very heavily.[19]

Two aspects of our setting allow us to identify utility functions separately from the discount factor. First, we have a finite horizon setting, where at the end of the year, the quotas are reset and all agents start with a fresh quota for the following year. This means that every December, the agent faces a static optimization problem, conditional on the sales agents's state ($\%AQ$). Utility parameters are well identified for a static model, and hence the agent's choice in the last period should allow us to non-parametrically identify the agent's utility function. Given this, variation in sales (that is monotonically linked to effort) in the last period and variations in wealth should help identify the effort disutility and risk aversion coefficient within the utility function.

Second, the bonus setting generates exclusion restrictions between current and future utility; i.e., we have instruments in non-bonus periods that do not affect current utility, but only future utility. As we demonstrated with reduced form evidence earlier, the fact that an agent's performance in November is related to his proximity to the annual bonus that will only be given in December indicates forward looking behavior. We also demonstrated other evidence of how agents respond to quarterly or annual quotas even though they do not affect current payoffs. This allows us to estimate discount factors.

Third, the ability to identify both parameters of a hyperbolic discount function is due to a combination of the two aspects above. As shown in the numerical example in section 3.2, given a utility function, exponential discounters and hyperbolic discounters have different sequences in their inter-temporal effort and sales. Specifically, the exponential discounter's effort is smoother and more evenly spread across the pre-bonus and bonus periods compared to the hyperbolic

[19] Another reason why an agent might exert low effort is that they may have wrong expectations about the transition density of future states, i.e., they may be very pessimistic about future good states. Like other dynamic structural modeling papers, we assume rational expectations for the transition densities of states. In this case, this translates into a rational expectations assumption on sales shocks.

discounter who puts in substantially less effort in the non-bonus period and more effort in the bonus period. Since the utility function is entirely identified off the last period choices and we have exclusion restrictions for identifying forward looking behavior, we can take advantage of the differences in inter-temporal sequences of effort (sales) predicted by the two models to identify the present bias factor in addition to the long term discount factor.

Beyond these conceptual arguments, we designed a simulation to illustrate that discount factors and utility parameters can be identified empirically in this setting. We consider the following model for the simulation. Let realized sales be a function of effort and normal random sales shocks $y_{it} = e_{it} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0,\sigma^2)$. Let salesperson utility be given $u_{it} = -de_{it}^2 + E[W_{it}]$, where wealth is $W_{it} = BI_{t=T}I_{\{s_{it}+y_{it}>Q\}}$, with B the bonus for achieving quota Q and the cumulative sales state evolving as follows: $s_{it+1} = s_{it} + y_{it}$. To keep the setting simple, we consider a three period model, where the bonus is paid at the end of the third period, i.e., T=3. Further, we set B=60 and Q=30.

We varied the simulated number of individuals from 50 to 1000. The true values and the estimates and standard errors for each simulation are reported in Table 5. The disutility parameters, discount factors and the standard deviation of the sales shocks are all estimated very precisely, lending confidence to the identification arguments above.

6 Results

We first report the first stage estimates of the demand shifters and effort policy function for the sales response model; then we report estimates of structural parameters of sales agents' utility functions from the second stage estimation. In the second stage estimation, we perform a grid search over the discount parameters (β and δ). We then perform several counterfactual simulations to address the substantive questions we seek to answer.

6.1 First Stage Estimates

The parameter estimates for the demand shifters in the sales response function is reported in Table 6a. We find that only the interaction term between lagged annual quota and indirect sales revenue are statistically significant. Thus larger markets tend to have a bigger sales multiplier independent of effort in high demand periods.

We estimate segment level effort policy functions by estimating the non-parametric relationship between sales and state variables through Chebyshev polynomials of the state variables. We estimated up to fourth order Chebyshev polynomials with alternative number of segments and choose the best fitting model based on the Bayesian Information Criterion (BIC). Uniformly, three segment models had the best fit. The estimates of the best fitting polynomial function and the standard deviations of the revenue shocks for each segment are reported in Tables 6b and 6c. As the coefficients associated with the Chebyshev polynomials have no intuitive meaning, for intuition, we show graphs of the effort policy function for the three segments as a function of percentage annual quota ($\%AQ$) for select months in Figure 6a. $\%AQ$ is normalized across sales agents, such that 1 implies at quota and 0.9 indicates 10% below quota and 1.1 indicates 10% above quota.

Table 7 shows the share of the three segments and their descriptive characteristics. Segment 2 is the largest with a share of 47%; Segments 1 and 3 have shares of 32% and 21% respectively. The average tenure with the firm is not very different across segments at approximately 12 years. Segment 3 has the highest annual quotas, followed by Segment 2 and Segment 1. Interestingly, Segments 2 and 3 with larger quotas achieve their quota targets more often than Segment 1 which has trouble meeting quota.

Figure 6a shows the Segment 3 exerts the most effort and is the most productive segment, and Segment 1 exerts the least effort and is the least productive segment. This is consistent with the allocated quotas and percentage of time quotas are achieved in Table 7. We also see a positive relationship between exerted effort and $\%AQ$ for all months shown. As for $\%QQ$, we see an increasing but concave relationship in March implying that once a sales person is way above the quarterly quota she starts to gradually slow down. Given that the average states in March for each segment were 0.55, 0.58, 0.62, respectively, not a lot of sales people are in the position to slow down. Effort in December does not fall off even if the sales person has already reached or exceed quota *(%AQ*>1), likely due to the overachievement commissions in preventing sales people from lowering effort after achieving quota. Our results are consistent with Steenburgh (2008), who finds that sales people "give up" when far away from achieving quota, such as for all segments in our case, but do not slow down much once quota is reached.

Figure 6b shows the effect of tenure on effort for all segments. Sales people in segment 2 and 3 initially increase effort with experience, but this tapers off with time. This is probably due to the fact that in the early years of their careers, they want to work hard not only for monetary payments from increased wages but also other intangible incentives such as promotions or transfers to better job titles. However, after a certain amount of years, these intangibles don't matter as much and the effort levels tend to taper off. Interestingly, Segment 1, the lowest productivity segment, does not gain in productivity from experience.

6.2 Second-Stage Structural Parameter Estimates

Discount Factor

We performed a grid search over the set of discount parameters in steps of 0.01 for δ and 0.1 for β. Table 8a presents the mean absolute percentage errors (MAPE) associated with each set of hyperbolic parameters where $\beta=1$ represents exponential discounting. A β of 0.8 and a δ of 0.92 has the lowest MAPE. Thus our estimates show a distinct present bias in that $\beta<1$.

Frederick, Loewenstein and O'Donoghue (2002) have a comprehensive summary of the estimated discount factors from previous studies. The summary shows that the estimated discount factors vary extensively ranging from as low as a mere 0.02 to no discounting at all with a discount factor of 1. For purely monetary values, the estimated discount factor seems rather low. But as Frederick, Loewenstein and O'Donoghue (2002) point out, for behavioral aspects such as pain and thus in our case effort, the discount factors tend to be low and hence our estimates appear reasonable.

Utility

The first column of Table 8b reports the structural parameter estimates of the sales agent's utility function for a forward looking sales person, consistent with the model we developed earlier. Overall, the disutility parameters for all three segments are negative and significant. These estimates are consistent with the effort policy functions estimated in the first stage. Segment 3, which produces the greatest sales on average, has the lowest disutility for effort. Segment 1, which has the lowest sales, has the greatest disutility. The risk aversion coefficients for all segments are insignificant showing no direct evidence of risk aversion by the sales agents. This may be because in the range of incomes earned by the sales force, risk aversion is not a serious concern. The estimated model fits the observed sales revenue data reasonably well with a MAPE of 10.7%.

6.3 Assessing the value of a dynamic structural model

How important is it to model the dynamics of salesperson behavior? In a static model, any effort would be attributed to current payoff, not accounting for the large future bonuses. This will downward bias the salesperson's disutility parameters and overstate the effects of compensation on productivity. Column 2 of Table 8b reports the estimates of the myopic model – discount factor set to zero. As expected, the disutility parameters are smaller in magnitude relative to the forward looking model for all segments. For Segment 3 the downward bias is as much as 22%. The myopic model also has a poorer fit: a MAPE of 18.8% relative to the MAPE of 10.7% for the dynamic model.

We next compare the revenue and effort predictions between the dynamic and myopic models. To isolate the effects of forward looking behavior, we simulate based on the disutility estimates from the dynamic model, but set the discount parameters to zero for the myopic model. Figure 7 compares the predicted revenues and effort of the myopic and dynamic models. The myopic agent has systematically lower revenues because she does not take into account the effect of future bonuses and overachievement commission in current effort. In contrast, the forward looking agent anticipates that in an uncertain environment, there is a chance of bad shocks later, which may prevent getting to the quota, so they prepare for such a rainy day by working harder early on so that they are within striking target of quota even if a bad sales shock occurred.

The effort graph in Figure 7 enables us to isolate out the sales revenue cyclicality and focus on the differences in effort across dynamic and myopic agents. The myopic salesperson concentrates much more effort in the bonus period, but the forward looking sales person smoothes effort over time, given the uncertainty in future demand shocks. The effort peaks in the bonus periods are not as pronounced for the dynamic consumer. The observed effort smoothing is similar to consumption smoothing by forward looking consumers facing uncertain incomes in the development economics literature.

6.4 Counterfactual Simulations

We now perform a series of counterfactual simulations that address the two sets of substantive questions we wish to answer. First, we address the issue of how valuable different components of the compensation plan are. The overall change in revenues under the alternative conditions is reported in Table 9 and the effect by segment in Table 10. Second, we compare the role of bonus frequency— how quarterly and annual bonuses affect performance.

Value of Quotas and Bonuses

We compare changes in revenues and profits when the firm moves from the current compensation plan to a pure commission-only plan. We consider two cases: (1) where the commission rate is the same as the current commission rate; and (2) a higher commission rate is such that total compensation is exactly equal to the current compensation. We find that the revenues are about 20.8% greater with the current compensation plan compared to a pure commission plan. Interestingly, we find from Table 10, all segments suffer from substantially poorer performance when quotas are removed and the firm shifts to a pure commission scheme. Even when adjusting commission rates to be higher to make total compensation identical to current levels, we find that revenues are about 4% higher.

Value of Overachievement Compensation

We compare changes in revenues and profits when the firm eliminates the over-achievement commission rate, which motivates sales people who are close to reaching their quota to continue exerting effort. Overall revenues drop by 13.3% and even accounting for the additional commission costs, profits are lower by about 2% (assuming gross margin of 33%).

Figure 8a plots the effort level of sales agents who met and didn't meet the annual quota, respectively. For those who met the annual quota, the effort level does not decline even when close to the quota because of the overachievement commission. In contrast, those who did not meet quota decrease effort towards the end of the year as they are unlikely to meet quota and therefore overachievement commission has no impact on their earnings. Thus overachievement commission provides the incentives for the most productive sales people even if they have already met quota (or likely to meet quota). Not surprisingly, Table 10 indicates that overachievement commissions have the most impact on Segments 2 and 3.

Value of Cumulative Annual Quota

Rather than have a cumulative annual quota, what would be the effect of replacing it with just a fourth quarter quota? To study this, we remove the overachievement commission (which is based on reaching the annual quota) and split the total bonus payments across all four quarters. Overall, revenues drop by 15.8%. This decrease is greater than the 13.3%, where we just dropped the overachievement commissions. Thus the cumulative annual quota induces sales agents to exert greater effort and raise revenues by 2.5%.

We also consider the case where we split the annual quota into a quarterly bonus and an annual bonus so that people do not "give up" in the last quarter when they are far away from quota. While this did increase the effort in the last quarter, it reduced revenues overall because sales people did not put in as much effort earlier in the year to be within striking distance of annual quota, because it is not as large. Total revenues drop by 1%.

From Table 10, we see that removing overachievement commissions have the greatest impact on the productivity of Segment 3, the most productive segment. Revenues drop by about 17% for this segment, while the effect on the least productive segment is substantially smaller at 7%.

Quota-Bonus Frequency

We next investigate the value of quarterly bonuses relative to annual bonuses. Figure 8b shows the comparison of effort between the current plan and when quarterly bonuses are eliminate and only the annual bonus is left. Effort drops consistently across the year when there are no quarterly quotas. Overall revenues fall by 5%. Even in December, when there is the annual bonus on the table, revenue falls by 2% and effort falls by 4%. Thus annual bonus and over-achievement commissions have less of an impact on year-end performance without quarterly bonuses. Why?

The quarterly bonus induces sales agents to work harder in a given quarter. But it also helps them achieve the annual quota by helping them stay on track of their annual goal. Without quarterly bonus, sales agents do not have much incentive to work hard early on. This lack of incentive leads them to be farther away from the annual quota by December. Annual bonuses and over-achievement commission have little impact on effort as sales agents are more likely to give up meeting annual quota.

The impact of quarterly bonuses also differs across the three segments of consumers. Table 10 indicates that quarterly bonuses have relatively minimal impact on Segment 3, the most productive segment, but very high impact on Segment 1. In effect, quarterly bonuses are needed as pacers to the less productive sales people than for the most productive sales people.

To the best of our knowledge, there has been no analysis to-date on what is the appropriate frequency of quota and bonuses. There has been some descriptive work in the education literature on how frequent testing affects academic performance (for an extensive survey, see Bangert-Drowns et al. 1991) and some experimental work in behavioral psychology (Heath, Larrick and Wu, 1999). The basic idea is that achieving short-term goals make achieving long-term goals more feasible. Our analysis show that the short term goals are more valuable to the least productive segment; i.e., in education terms it may imply that weaker students gain more by periodic testing, relative to stronger students who would study independent of exams.

7 Conclusion

Personal selling is a primary marketing mix tool for most B2B firms to generate sales, yet there is little research on how the compensation plan motivates the sale-force and affects performance. This paper develops and estimates a dynamic structural model of salesforce response to a compensation plan with various components: salary, commissions, lump-sum bonus for achieving quotas, and different commission rates beyond achieving quotas. Our analysis helps us assess the impact of (1) different components of compensation and (2) the differential importance of periodic bonuses on performance on different segments of sales people. We find that the quota-bonus scheme used by this firm increases performance of the sales force by serving as stretch goals and pushing employees to accomplish targets. Features such as overachievement compensation reduce the problems associated with sales agents slacking off when they get close to achieving their quota.

Further, quarterly bonuses serve as a continuous evaluation scheme to keep sales agents within striking targets of their annual quotas. In the absence of quarterly bonuses, failure in the early periods to accomplish targets caused agents to fall behind more often than in the presence of quarterly bonuses. Thus, quarterly bonus serves as a valuable sub-goal which helps the sales force stay on track in achieving their overall goal; they are especially valuable to low performers. In contrast, overachievement commissions increase performance among the highest performers.

We use recent innovations in the two-step dynamic structural model estimation to accommodate unobserved heterogeneity in sales force response. The approach is flexible, yet computationally feasible with minimal additional burden compared to traditional two-step methods. The bonus setting also allowed us to estimate discount factors for a hyperbolic discounting model, which has traditionally been impossible with naturally occurring data.

We now discuss limitations of the paper, which provide promising avenues for future research. First, effort tends to be multi-dimensional and one possibility is that quotas and bonuses force people to focus on the effort that lead to final sales in bonus periods, while agents may focus on earlier stages of the selling process in non-bonus periods. Such a multidimensional effort cannot be identified merely from sales data. We hope data from CRM databases which track customer stages through the selling process can help shed insight on this issue. We believe this is an exciting area for future research.

Second, compensation contracts can serve to select the right type of sales people. We do not address selection issues. One possibility is to use a longer panel of sales people's performance that includes attrition information. If there were variation in contracts that affected employee retention, that could also help address this problem. One needs more work on scenarios with richer contracts. For example, one could study peer effects on sales performance and selection effects when firms shift from individual to team based compensation (Chan, Li and Pierce 2009).

In summary, this paper provides a rigorous framework to empirically understand how the sales force responds to a very rich compensation structure involving many components of compensation: salary, commissions, quota and bonuses at quarterly and annual frequencies. Our analysis helps obtain a number of useful substantive insights. Nevertheless, the issues raised above provide an interesting agenda for future work.

References

Albers, Sonke, Murali K. Mantrala, and Shrihari Sridhar (2010), "Personal Selling Elasticities: A Meta-Analysis," *Journal of Marketing Research*, 47-5, pp. 840-853

Ainslie, G. (1992), "Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person," *Cambridge University Press*.

Arcidiacono, Peter and John B. Jones (2003), "Finite Mixture Distributions, Sequential Likelihood, and the EM Algorithm," *Econometrica,* 71-3, pp. 933-946.

Arcidiacono, Peter and Robert A. Miller (2011), "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Econometrica*, forthcoming.

Assmus, Gert, John U. Farley and Donald R. Lehmann (1984), "How Advertising Affects Sales: Meta-Analysis of Econometric Results," *Journal of Marketing Research*, 21-1, pp. 65-74.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin (2007), "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75-5, pp. 1331-1370.

Bangert-Drowns, Robert L., James A. Kulik and Chen-Lin C. Kulik (1991), "Effects of frequent classroom testing," *Journal of Educational Research*, 85-2, pp. 89-99.

Basu, Amyia K., Rajiv Lal, V. Srinivasan and Richard Staelin (1985), "Sales-force Compensation Plans: An Agency Theoretic Perspective," *Marketing Science*, 8-3, pp. 324-342.

Beauchamp, Andrew (2010), "Abortion Supplier Dynamics," Working Paper, Duke University.

Bijmolt, Tammo H.A., Harald J. Van Heerde and Rik G.M. Pieters (2005), "New Empirical Generalizations on the Determinants of Price Elasticity," *Journal of Marketing Research* 42-2, pp. 141–56.

Chan, Tat Y., Jia Li, and Lamar Pierce (2009), "Competition and Peer Effects in Competing Sales Teams," Working Paper, Washington University, St. Louis.

Chevalier, Judith and Austan Goolsbee (2009), "Are Durable Goods Consumers Forward-Looking? Evidence from College Textbooks," *The Quarterly Journal of Economics*, 124-4, pp. 1853-1884,

Copeland, Adam and Cyril Monnet (2009), "The Welfare Effects of Incentive Schemes," *Review of Economic Studies*, 76-1, pp. 96-113.

Darmon, Rene (1997), "Selecting Appropriate Sales Quota Plan Structures and Quota-setting Procedures," *Journal of Personal Selling & Sales Management*, 17-1, pp. 1-16.

Elster, Jon (1979), "Ulysses and Sirens: Studies in Rationality and Irrationality" *Cambridge University Press.*.

Ferrall, Christopher, and Bruce Shearer (1999), "Incentives and Transactions Costs within the Firm: Estimating an Agency Model Using Payroll Records," *Review of Economic Studies*, 66-2, pp. 309-338.

Finger, Stephen R. (2008), "Research and Development Competition in the Chemicals Industry," Working Paper, Duke University.

Frederick, Shane, George Loewenstein and Ted O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review", *Journal of Economic Literature*, 40-2, pp. 351-401.

Heath, Chip, Richard P. Larrick and George Wu (1999), "Goals as Reference Points," *Cognitive Psychology*, 38, pp. 79-109.

Heide, Christen P. (1999), "Dartnell's 30th Sales Force Compensation Survey," Dartnell Corporation; 30th edition.

Holmstrom, Bengt (1979), "Moral Hazard and Observability," *Bell Journal of Economics*, 10-1, pp. 74-91.

Holmstrom, Bengt and Paul Milgrom (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55-2, pp. 303-328.

Hotz, V. Joseph. and Robert A. Miller (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60-3, pp. 497-529.

Jiang, Renna and Robert W. Palmatier (2009) "Structural Estimation of a Moral Hazard Model: An Application to Business Selling," Working Paper, UC Davis.

Joseph, Kissan and Manohar Kalwani (1998), "The Role of Bonus Pay in Sales force Compensation Plans," *Industrial Marketing Management*, 27-2, pp. 147-159.

Judd, Kenneth L. (1998), "Numerical Methods in Economics", MIT Press.

Kim, Son Ku (1997), "Limited Liability and Bonus Contracts," *Journal of Economics and Management Strategy*, 6-4, pp. 899-913.

Laibson, David (1997), "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112-2, pp. 443-477.

Laibson, David (1998), "Life-Cycle Consumption and Hyperbolic Discount Functions," *European Economic Review*, 42, pp. 861-871.

Lal, Rajiv and V. Srinivasan (1993), "Compensation Plans for Single-and Multi-product Sales forces: An Application of the Holmstrom – Milgrom Model," *Management Science*, 39-7, pp. 777-793.

Magnac, Thierry and David Thesmar (2002), "Identifying dynamic discrete choice models," *Econometrica*, 70-2, pp. 801-816.

Misra, Sanjog and Harikesh Nair (2011), "A Structural Model of Sales-force Compensation Dynamics: Estimation and Field Implementation," *Quantitative Marketing and Economics*, forthcoming.

Oyer, Paul (1998), "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality," *Quarterly Journal of Economics*, 113, pp. 149-185.

Oyer, Paul (2000), "A Theory of Sales Quotas with Limited Liability and Rent Sharing," *Journal of Labor Economics*, 18, pp. 405-426.

Paarsch, Harry, J. and Bruce Shearer (2000), "Piece Rates, Fixed Wages, and Incentive Effects: Statistical Evidence from Payroll Records," *International Economic Review*, 40, pp. 59-92.

Park, Eun-Soo (1995), "Incentive Contracting under Limited Liability," *Journal of Economics and Management Strategy*, 4, pp. 477-490.

Phelps, Edmund. S. and Robert Pollak (1968), "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35, pp. 185-199.

Raju, Jagmohan S., and V. Srinivasan (1996), "Quota-Based Compensation Plans for Multi-Territory Heterogeneous Sales-Forces," *Management Science*, 42, pp. 1454-1462.

Rao, Ram C. (1990), "Compensating Heterogeneous Sales-forces: Some Explicit Solutions", *Marketing Science*, 9, pp. 319-342.

Rust, John 1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, pp. 999-1033.

Rust, John (1994), "Structural Estimation of Markov Decision Processes," in Handbook of Econometrics, Volume 4, ed. by R.E. Engle and D. McFadden, Amsterdam: Elsevier-North Holland, Chapter 14, pp. 3082-3139.

Samuelson, Paul (1937), "A Note on Measurement of Utility," *Review of Economic Studies*, 4, pp. 155-161.

Steenburgh, Thomas (2008), "Effort or Timing: The Effect of Lump-Sum Bonuses," *Quantitative Marketing and Economics*, 6, pp. 235-256.

Sudhir, K. (2001), "Structural Analysis of Manufacturer Pricing in the Presence of Strategic Retailer," *Marketing Science*, 20-3, pp. 244-264.

Thaler, Richard H. (1982), "Some Empirical Evidence on Dynamic Inconsistency," *Economic Letters*, 8, pp. 201-206.

Zoltners, Andris A., Prabhakant Sinha and Sally E. Lorimer (2008), "Sales Force Effectiveness: A Framework for Researchers and Practitioners," *Journal of Personal Selling & Sales Management*, Special Issue on Enhancing Sales Force Productivity, 28, pp. 115-131.

Table 1: Firm's Compensation Plan

Type	Description	Payment period
Quarterly Bonus	$1500 Awarded if quarterly revenue exceeds quarterly quota	Mar, Jun, Sep
Annual Bonus	$4000 Awarded if annual revenue exceeds annual quota	Dec
Base Commission	About 1.5%* paid in proportion to the revenue generated each month	Every month
Overachievement Commission	About 3%* paid in proportion to the total cumulative revenue surpassing the annual quota	Dec

**These numbers are approximate for confidentiality reasons.*

Table 2: Descriptive statistics

Salespeople	348				
Average salary (USD)	$3,585				
Average tenure	11.8				
	Average Quota '000 USD	% Achieving Quota	Average Sales '000 USD	Indirect Sales (ISR) Index	Average Quota / ISR Index
Quarter 1	232.4	51.1	273.0	3.1	75.2
Quarter 2	374.2	49.8	445.0	5.3	71.0
Quarter 3	397.1	42.8	407.0	4.5	88.2
Quarter 4	-	-	565.6	6.6	
Annual	1639.3	49.9	1690.6	19.5	84.3

Table 3: Sales Performance in November

	$\%AQ$<0.5	$\%AQ$>0.5
	Estimate	Estimate
Intercept	0.05*** (0.018)	0.06*** (0.0123)
$\%AQ$	0.06 (0.049)	0.04*** (-.0126)

*** $p < 0.01$

Table 4: Testing for Sales Substitution across Months

	Model 1	Model 2	Model 3
Last Month Sales	0.302*** (0.014)	0.188*** (0.013)	0.188*** (0.014)
Qtr 1st Month *Last Month Sales	-0.007 (0.013)	0.023 (0.015)	
Qtr 1st Month*Last Month Sales* "Way off Target Last Qtr"			0.029 (0.018)
Qtr 1st Month*Last Month Sales* "Not Way off Target"			0.004 (0.024)
Monthly Allocated Quota		0.565*** (0.021)	0.566*** (0.021)
Indirect Sales		14.063*** (2.130)	14.15*** (2.229)
Sales person Fixed Effects	Yes	Yes	Yes

***: p<0.01

Table 5: Simulation Results

	True Values			
# Individuals	*d=-0.1*	β*=0.6*	δ*=0.95*	σ*=5*
50	-0.105	0.678	1.035	5.686
	(0.002)	(0.093)	(0.053)	(0.368)
100	-0.100	0.607	0.938	4.966
	(0.002)	(0.040)	(0.020)	(0.181)
200	-0.101	0.612	0.937	4.946
	(0.002)	(0.024)	(0.015)	(0.101)
500	-0.100	0.619	0.918	4.844
	(0.002)	(0.013)	(0.014)	(0.111)
1000	-0.100	0.601	0.944	4.938
	(0.001)	(0.009)	(0.013)	(0.056)

Table 6a: Parameter Estimates – Sales Response

Lagged annual quota	0.002
	(0.005)
Indirect sales	-6.735
	(5.554)
Indirect sales*Lagged annual quota	0.022***
	(0.003)

***: $p<0.01$

Table 6b: Parameter Estimates – Effort Policy Function

	Segment 1	Segment 2	Segment 3
ρ_0	-18.63	**-74.23****	**-263.00*****
	(26.09)	(32.74)	(89.28)
$\rho_1(z_1)$	**141.10*****	**146.42*****	**312.36****
	(44.32)	(55.41)	(137.85)
$\rho_2(z_1)$	-23.54	-21.67	**-148.39*****
	(17.09)	(22.37)	(63.23)
$\rho_3(z_1)$	6.41	6.98	15.87
	(3.94)	(5.14)	(13.09)
$\rho_1(z_2)$	**102.26*****	**241.73*****	**458.59*****
	(38.00)	(44.93)	(117.20)
$\rho_2(z_2)$	-22.13	**-62.91*****	-87.28
	(18.67)	(21.89)	(55.50)
$\rho_3(z_3)$	**8.09***	**11.34*****	15.90
	(4.39)	(4.54)	(10.79)
$\rho_1(z_1)\rho_1(z_2)$	-114.52	-31.98	110.00
	(75.28)	(82.95)	(181.61)
$\rho_1(z_1)\rho_2(z_2)$	-16.60	0.27	**-113.83*****
	(19.77)	(18.40)	(40.33)
$\rho_2(z_1)\rho_1(z_2)$	9.07	-10.14	**86.81*****
	(14.63)	(17.08)	(36.85)
$\rho_1(z_1)\rho_1(IS)$	**-20.00****	8.12	42.42
	(8.95)	(11.31)	(29.62)
$\rho_1(z_2)\rho_1(IS)$	-1.11	**16.76*****	**20.10*****
	(2.68)	(3.18)	(8.53)
$\rho_1(z_1)\rho_1(z_2)\rho_1(IS)$	9.52	**-43.91*****	**-89.11****
	(17.62)	(18.88)	(44.87)
$\rho_1(z_1)\rho_1(AQ_{y-1})$	**-0.04****	**-0.03****	**-0.07*****
	(0.02)	(0.02)	(0.03)
$\rho_1(z_2)\rho_1(AQ_{y-1})$	-0.01	**-0.06*****	-0.02
	(0.01)	(0.01)	(0.01)
$\rho_1(z_1)\rho_1(z_2)\rho_1(AQ_{y-1})$	0.04	**0.05*****	0.02
	(0.04)	(0.03)	(0.04)
$\rho_1(\tau)$	-0.53	**9.00*****	**18.53*****
	(1.38)	(2.14)	(6.90)
$\rho_2(\tau)$	0.05	**-0.27*****	**-0.52****
	(0.05)	(0.07)	(0.26)
$\rho_3(\tau)$	0.001	**0.002*****	0.004
	(0.000)	(0.001)	(0.003)

***: p<0.01, **: p<0.05, *: p<0.1

Table 6c: Revenue Shock Distribution – Standard Deviation

	Segment 1	Segment 2	Segment 3
Sigma	81.61	143.72	279.35

Table 7: Descriptive Characteristics of Segments

	Segment1	Segment2	Segment3
Share	0.32	0.47	0.21
Tenure*	11.5	12.2	11.5
Achieve quarterly quota - Q1	0.46	0.54	0.58
Achieve quarterly quota - Q2	0.38	0.55	0.62
Achieve quarterly quota - Q3	0.31	0.49	0.53
Achieve annual quota	0.30	0.57	0.64
Average annual quota**	1,201.4	1,615.7	2,363.7
Average December revenue**	130.2	273.0	559.1

*Tenure is measured in years **Average quotas and revenues are indicated in USD(K)

Table 8a: Optimal Discount Factor – Model Fit
Mean Absolute Percentage Error by Discount Factors

		δ									
		0.9	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99
β	0.4	0.1336	0.1347	0.1367	0.1378	0.1391	0.1382	0.1402	0.1428	0.1450	0.1475
	0.5	0.1249	0.1258	0.1243	0.1270	0.1255	0.1300	0.1266	0.1300	0.1287	0.1318
	0.6	0.1167	0.1181	0.1193	0.1203	0.1219	0.1200	0.1207	0.1167	0.1198	0.1241
	0.7	0.1103	0.1149	0.1134	0.1124	0.1107	0.1127	0.1140	0.1202	0.1220	0.1242
	0.8	0.1104	0.1084	0.1074	0.1098	0.1100	0.1132	0.1149	0.1150	0.1175	0.1172
	0.9	0.1121	0.1108	0.1101	0.1117	0.1148	0.1145	0.1180	0.1185	0.1192	0.1203
	1	0.1109	0.1098	0.1103	0.1168	0.1181	0.1182	0.1167	0.1223	0.1251	0.1328

Table 8b: Utility Parameters

	With Forward Looking	Without Forward
Segment 1		
Disutility	-0.240	-0.199
	(0.005)	(0.006)
Risk Aversion	-0.0001	-0.0001
	(0.0020)	(0.0003)
Segment 2		
Disutility	-0.119	-0.092
	(0.038)	(0.005)
Risk Aversion	-0.0001	-0.0001
	(0.0005)	(0.0005)
Segment 3		
Disutility	-0.061	-0.048
	(0.002)	(0.002)
Risk Aversion	0.0000	0.0000
	(0.0002)	(0.0004)

***p<0.01

Table 9: Impact of Alternative Bonus Plans on Sales Revenues

Counterfactual	**Change in Revenues**
1a. Only Pure Commissions	-20.8%
1b. Only Pure Commissions (adjusted to equal payout with bonus)	-3.8%
2a. No Bonus (Only Commissions + Overachievement Commission)	-9.3%
2b. No Bonus (Commissions adjusted to equal payout with bonus)	-1.5%
3. No overachievement commissions	-13.3%
4a. Cumulative Annual Quota replaced with quarterly quota	-4.2%
4b. Annual Bonus split into Quarterly and Annual Bonus	-1.0%
5a. Remove quarterly bonus	-4.6%

Table 10: Impact of Alternative Bonus Plans on Sales Revenues by Segment

% decrease from different components	Seg1	Seg2	Seg3
Pure commission	17.9%	21.0%	21.4%
Without overachievement	7.0%	12.6%	17.1%
Without quarterly bonus	10.0%	4.5%	2.0%

Figure 1: Types of Incentive Compensation Schemes

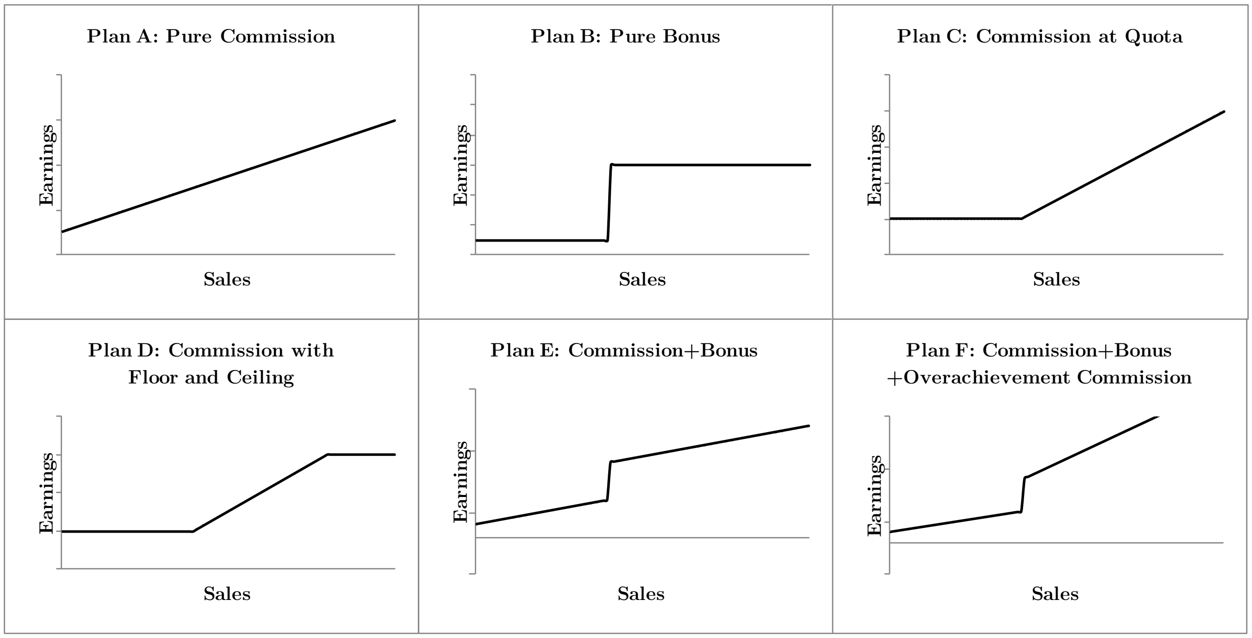


Figure 2a: How Quotas and Bonus Serve as Stretch Goals

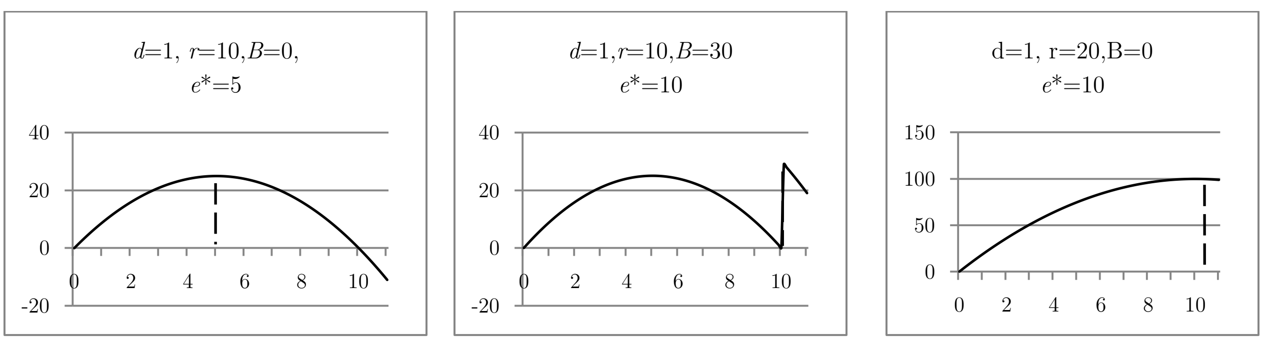


Figure 2b: Effort as a Function of Distance to Quotas

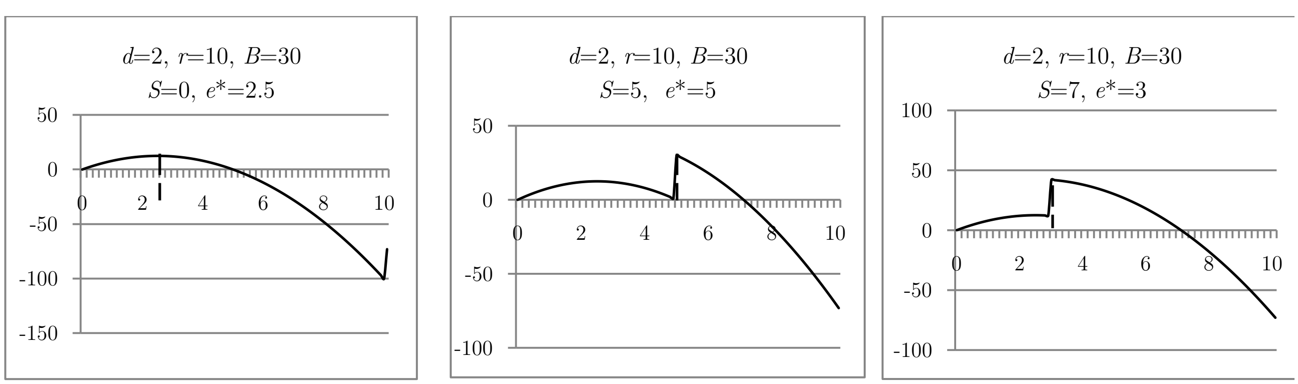


Figure 2c: Difference in Behavior between Hyperbolic and Exponential Discounters

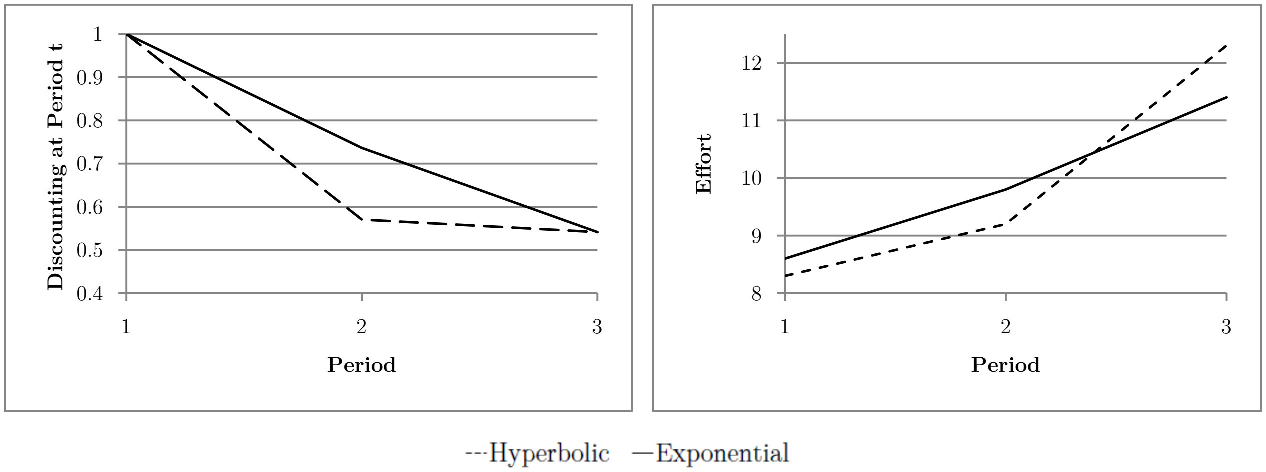


Figure 3: Fraction of People Achieving Quota in December

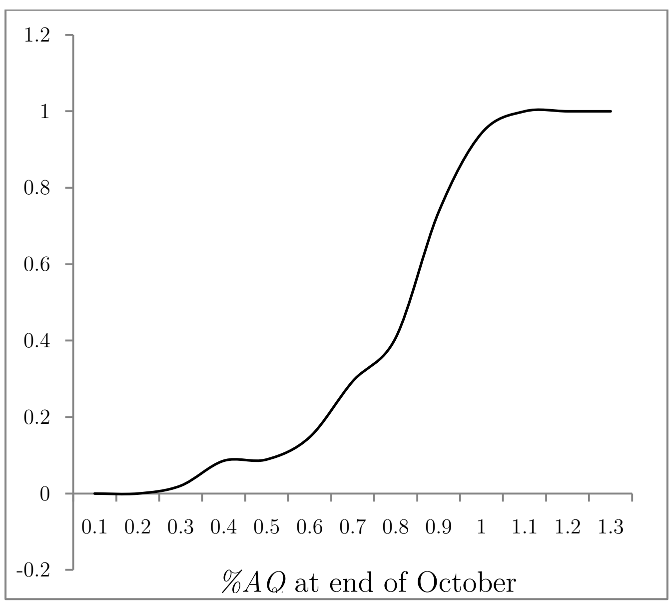


Figure 4a: Sales and Percentage Quota Achieved – Bonus Months

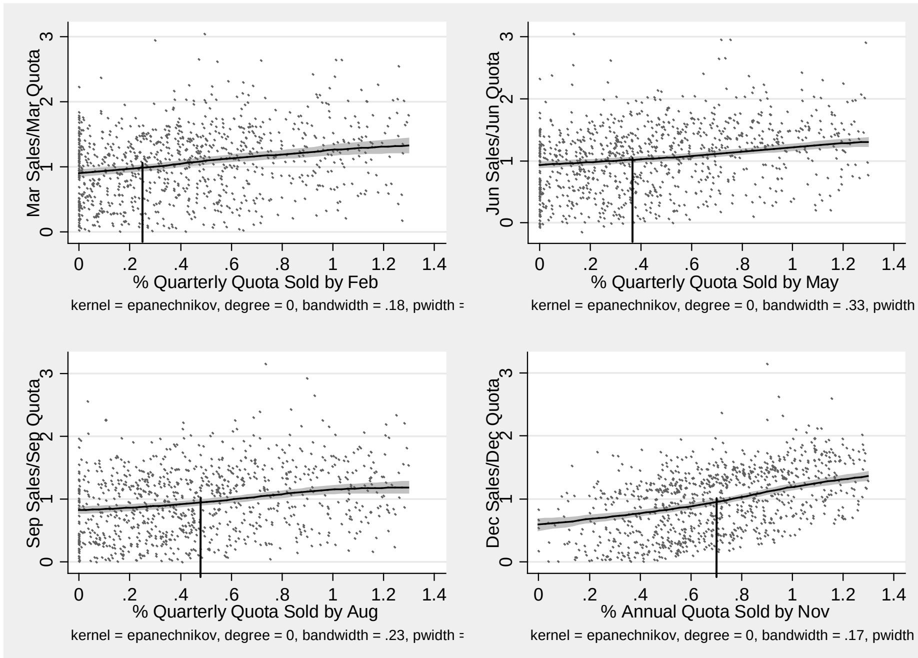


Figure 4b: Sales and Percentage Quota Achieved – Pre-Bonus Months

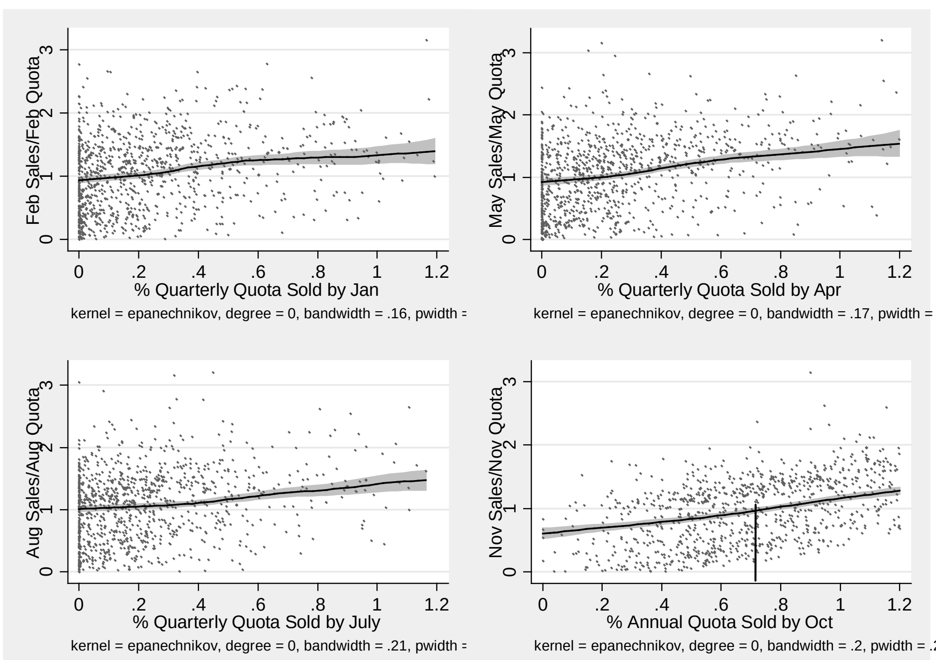


Fig 5a: Revenues from Regular Sales Force

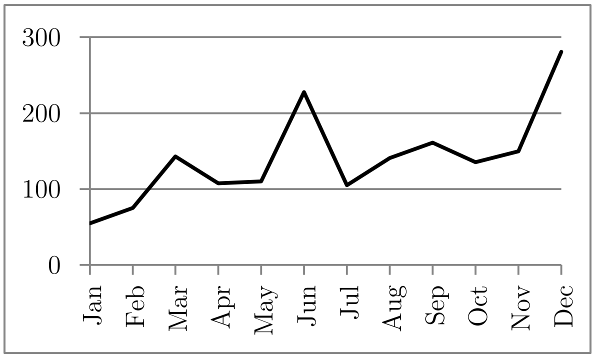


Fig 5b: Indirect Sales Revenue (ISR) Index[20]

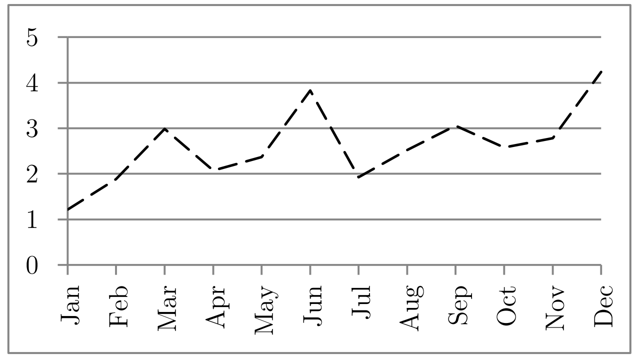
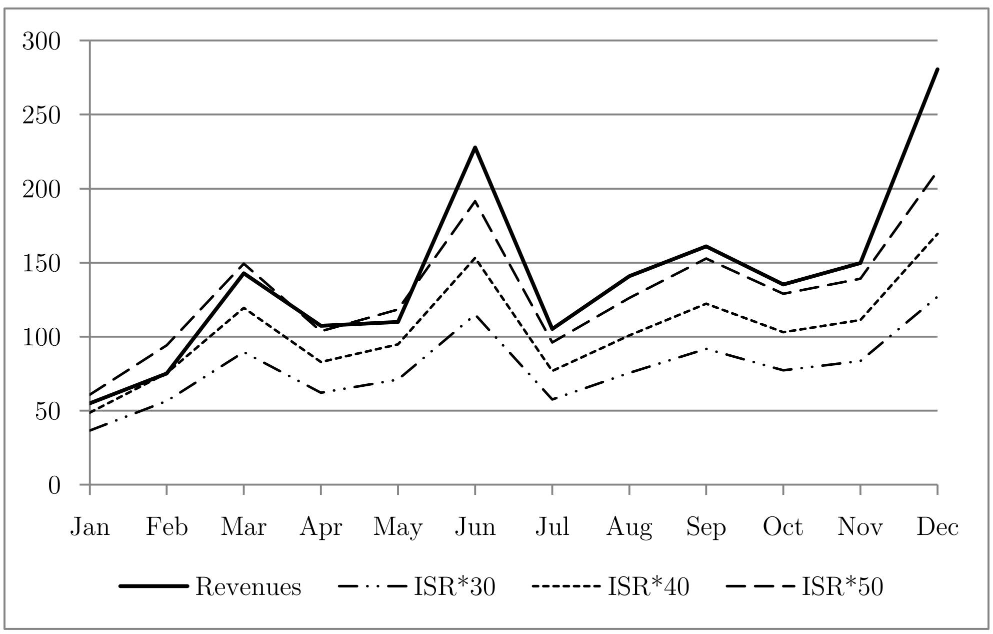


Fig 5c: Revenues from Regular Sales Force and ISR Index Multiples



[20] The indirect sales force revenue (ISR) index is on a base of 1 reflecting the indirect sales in January 1999. Here we have averaged the index across the three years from 1999-2001.

Figure 6a: Effort Policy by Segment as a Function of % Quota

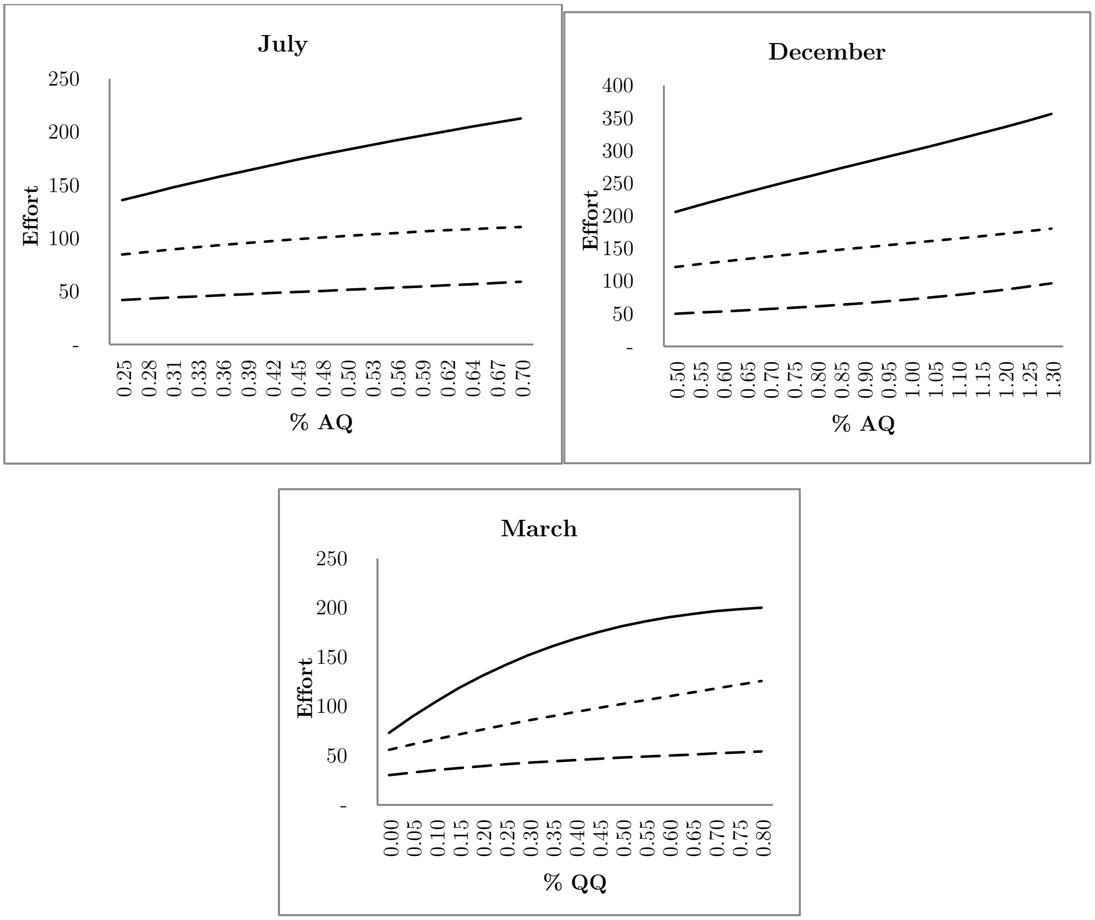


Figure 6b: The Effect of Tenure on Effort

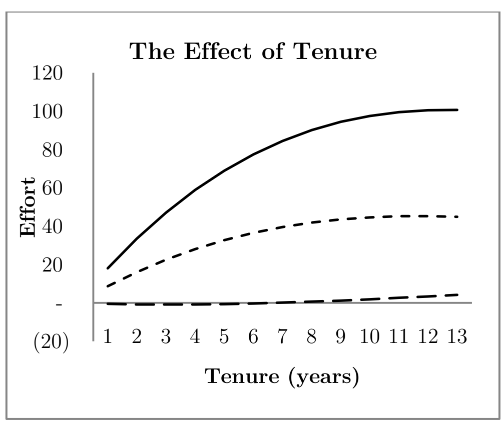


Figure 7: Simulated Revenue & Effort– Static vs. Dynamic

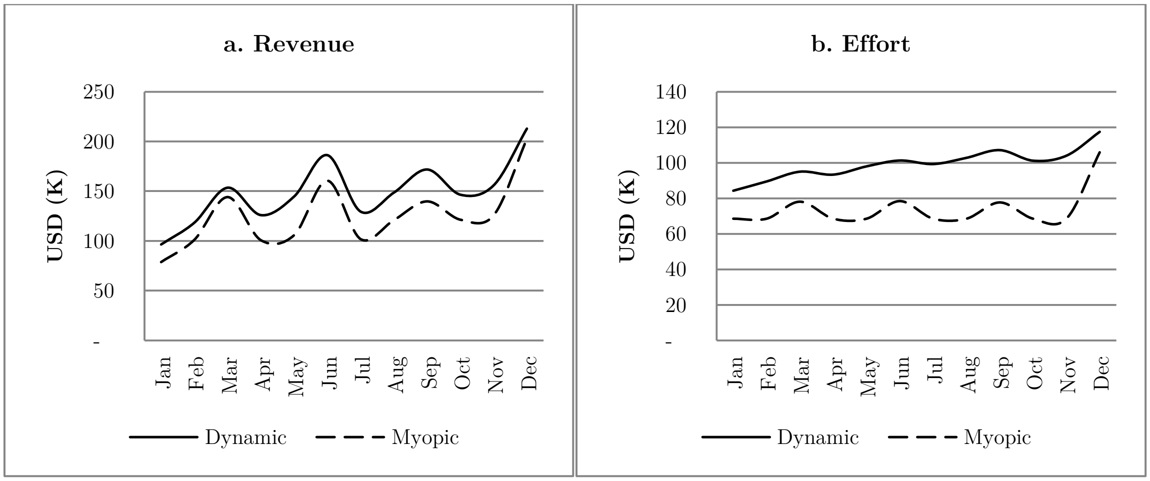


Figure 8a: Overachievement Commission

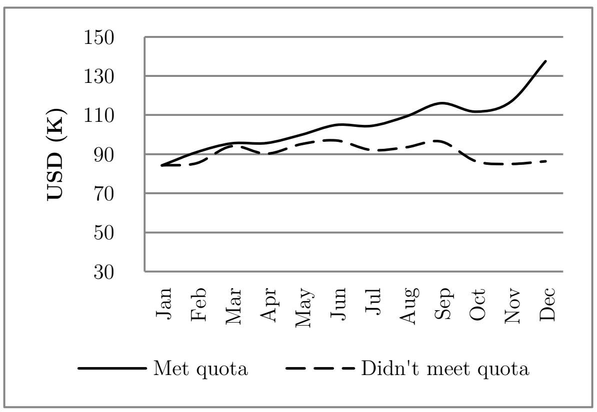


Figure 8b: Quarterly Quotas