Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Salesforce Contracting Under Demand Censorship

Leon Yang Chu, Guoming Lai,

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 15, No. 2, Spring 2013, pp. 320–334
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)
http://dx.doi.org/10.1287/msom.1120.0424

Salesforce Contracting Under Demand Censorship

Leon Yang Chu
Marshall School of Business, University of Southern California, Los Angeles, California 90089,
leonyzhu@usc.edu

Guoming Lai
McCombs School of Business, University of Texas at Austin, Austin, Texas 78712,
guoming.lai@mccombs.utexas.edu

We study salesforce contracting in an environment where excess demand results in lost sales and the demand information is censored by the inventory level. In our model, a firm contracts with a risk-neutral sales agent with limited liability whose effort increases the demand stochastically. The firm designs the incentive contract and invests in inventory; the agent decides the sales effort. We find that the sales-quota-based bonus contract is optimal in such an environment, and the quota should be set equal to the inventory level when the first-best solution is not attainable. We further reveal that demand censorship can introduce peculiar effects on the optimal sales effort and service level that the firm implements. From our analysis of the additive and multiplicative effort cases, we find that in the additive effort case, it can be optimal, under demand censorship, for the firm to induce an effort and maintain a service level both greater than those under the first-best solution. Scenarios also exist where the firm should induce zero effort. For the multiplicative effort case, the optimal sales effort under demand censorship is lower than the first-best effort, whereas the optimal service level is higher than the first-best service level. The agent earns zero rent in the additive effort case but may earn a positive rent in the multiplicative effort case. Finally, our numerical analysis shows that demand censorship can have a significant negative impact on the value of contracting with the sales agent, especially when the sales margin is low and the market uncertainty is high.

Key words: salesforce contracting; demand censorship; newsvendor
History: Received: November 15, 2011; accepted: October 29, 2012. Published online in *Articles in Advance* January 28, 2013.

1. Introduction

Firms rely on salesforce to promote demand. In the United States, the average expenditure on salesforce costs ranges typically from 10% to 40% of firms' sales revenues (Albers and Mantrala 2008); in total, U.S. firms are estimated to spend more than $800 billion on salesforce costs (Zoltners et al. 2008). Designing appropriate salesforce compensation is of importance to incentivize the salesforce and utilize the best of its effort.

The existing literature has explored salesforce contracting in various scenarios; however, it often neglects the impact of limited inventory (capacity) on the fulfillment of demand. According to Gruen et al. (2002), out-of-stocks remains a large problem for retailers, distributors, and manufacturers, with an average frequency of stockouts of about 8.3% in the worldwide consumer goods industry. Facing stock-outs or delay of shipment, customers will choose to buy from other places, buy other brands, delay purchases, or simply cancel orders, as revealed from the market studies (see, e.g., Schary and Christopher 1979, Emmelhainz et al. 1991, Anderson et al. 2006). In such scenarios, the true demand information is generally censored. Nevertheless, most of the existing literature on salesforce contracting assumes unlimited inventory (capacity) or perfect backordering, and uses demand as the instrument for contracting. Such an assumption can decouple saleforce incentive design from inventory planning and make both problems more analytically tractable, which, however, does not necessarily mirror the reality. On the other hand, companies in practice usually contract their salesforce based on their sales revenues with bonus or commission incentives (see, e.g., Joseph and Kalwani 1998, Oyer 2000, Johnson 2010) without explicitly considering the possibility of out-of-stocks. As a result, given that the incentive plan is tied to the sales, the sales and marketing personnel are often concerned about the availability of the products when they exert efforts to promote demand. In some situations, they will try to persuade the company to install some "ideal" inventory level in their favor, as learned from our conversations with promotion planning practitioners. Although it is widely acknowledged that integrating sales and operations planning is important, the impact of such a demand censoring issue driven by limited inventory on salesforce contracting was only questioned by Chen (2005) and remains unexplored in the academic literature.

In this paper, we study the integrated salesforce incentive design and inventory planning, taking demand censorship into account. Specifically, the firm in our model designs the salesforce contract and makes the inventory decision, based on which a risk-neutral agent that has limited liability decides the sales effort. Because of demand censorship, the salesforce contract has to condition on the realized sales instead of the demand. We find that in such an environment, the sales-quota-based bonus contract (i.e., the agent receives a bonus when the realized sales meet or exceed the quota and nothing otherwise) dominates other types of contracts. Demand censorship can also affect the implementability of the first-best solution. There are scenarios where the first-best solution could otherwise be achieved if the demand was not censored. In such scenarios, to achieve the best performance, the sales quota should be set equal to the inventory level.

Although it is typical in the salesforce literature that the second-best effort level is lower than the first best, demand censorship can introduce peculiar effects, as revealed from our analysis with two representative demand–effort relationships, the additive and multiplicative cases. We find that under demand censorship, it can be optimal for the firm to induce a sales effort greater than the first-best effort and maintain a service level higher than the first-best service level in the additive effort case, and there are also scenarios where it is optimal for the firm to induce zero effort even when the marginal cost for the agent to supply effort is negligible. For the multiplicative effort case under demand censorship, the optimal effort to induce is no higher than the first-best effort, but the optimal service level to maintain still exceeds the first-best service level. Hence, demand censorship can influence not only salesforce contracting but also the operations metric, the service level; furthermore, the effects of demand censorship can differ fundamentally for different demand–effort relationships. Finally, our numerical analysis shows that the effect of demand censorship on the value of contracting with the sales agent is the most significant when the margin of the sales is low and the market uncertainty is high. These results can be useful for firms to design efficient salesforce incentives and coordinate their inventory decisions in certain environments.

The remainder of our paper is organized as follows. Section 2 reviews the literature, and in §3 we describe the model. We present the benchmark in §4, and investigate the effects of demand censorship on the firm's salesforce contracting, inventory decision, and performance in §5. Three extensions of our model are discussed in §6, and we conclude in §7.

2. Literature Review

The salesforce contracting problem has attracted wide attention in the academic literature. Basu et al. (1985) studied the moral-hazard problem in the salesforce contracting context and revealed nonlinear shapes of the optimal compensation scheme with risk-averse salespeople. Lal and Staelin (1986) incorporated heterogeneity and asymmetric information about salespeoples' risk attitudes into the salesforce contracting problem, which interprets the rationale to design a menu of compensation schemes. The nonlinearity format of the optimal contracts derived from these research studies are, however, complex and might be difficult to apply in practice. With such a consideration, the subsequent studies by Lal and Srinivasan (1993) and Raju and Srinivasan (1996) revealed that the simple commission and bonus contracts might not lose much of the optimality in certain environments. Rao (1990) also showed that a commission plus bonus contract contingent on a quota can achieve optimality for motivating heterogeneous but risk-neutral salespeople.

Because quota-based bonuses are widely used in practice, Park (1995), Kim (1997), and Oyer (2000) aimed to develop theories to justify the use of such incentives. They focused on environments with risk-neutral salespeople who, however, assume limited liabilities. Park (1995) and Kim (1997) showed that the first-best solution can be achieved by a bonus contract if some regularity condition holds. However, the bonus contract is not the unique contract that can do so in their models. Focusing on this point, Oyer (2000) argued that the bonus contract can be more preferred than other compensation schemes in an additive effort environment if the hazard rate of the demand distribution is monotonically increasing and the agent's participation constraint is absent in the sense that there is no minimum expected compensation that has to be met (so that the firm can provide systematically less compensation as long as the limited liability constraint is ensured). Note that limited liability is sometimes assumed for both the agent and the principal in the literature. For example, Innes (1990) and Poblete and Spulber (2012) studied the effectiveness of debt instructions between the risk-neutral parties under double limited liability.

The above literature has neither studied any operational decision in conjunction with salesforce contracting, nor discussed the inventory or capacity constraint and demand censoring. Chen (2000, 2005) and Chen and Xiao (2009) studied salesforce incentives together with inventory decision. They considered settings where a firm needs to motivate the salespeople and elicit market information while making its inventory decision. Backorders are allowed, and thus the salesforce incentive contract can be contingent on the demand. They showed that commission or quota-commission incentives are optimal and reveal the importance to make the salesforce contracting and inventory decisions jointly. Jerath et al. (2010) investigated a setting where a sales manager exerts effort to increase demand and an operations manager simultaneously decides the inventory level and exerts effort to maintain the reliability of inventory supply. They explored compensation schemes that coordinate the decisions of the two managers. Hopp et al. (2010), Khanjari et al. (2012), and Chen and Xiao (2012) investigated joint supply chain contract and salesforce incentives with either a retailer- or manufacturer-employed salesperson. However, none of these studies considered demand censorship. One exception is the work of Heese and Swaminathan (2010). They investigated inventory planning together with sales effort decision in the presence of demand censorship. However, in their study, the firm makes both decisions, and the agency issue related to salesforce incentives is not studied.

Besides the effect that we explore on salesforce contracting, demand censorship has been investigated in the inventory literature where the underlying demand distribution is assumed to be not perfectly known. Lariviere and Porteus (1999) showed that due to demand censorship, a firm that is to learn the underlying demand distribution may stock excess inventory to limit the frequency of stockout to better acquire information. Research works by Ding et al. (2002), Bensoussan et al. (2007), Lu et al. (2006), and Chen and Plambeck (2008) have extended the investigation of inventory policies for perishable and nonperishable products under demand censorship with general distributions and Markovian demand processes; the "stock-more" result has been generally obtained. The "stock-more" result holds also in our study, but the driver originates from the objective to better incentivize the salespeople by limiting lost sales and to extract agency rent.

3. The Model

3.1. Problem Description

We consider a risk-neutral firm that sells a product through a sales agent. We focus on one selling season. Following the classical newsvendor literature, we assume that the inventory decision, q, is made before the selling season when the demand is still uncertain. Furthermore, we assume that if the realized demand exceeds the inventory level, the excess demand is lost. Such a setting is appropriate if the firm operates in a competitive environment with a long lead time, where substitutes of the product exist and customers are not willing to accept backorders. In contrast, if the demand is lower than the inventory level, without loss of generality, we assume that the excess inventory has zero value at the end of the selling season. Let c be the per-unit cost of obtaining the product, and let p ($> c$) be the per-unit selling price. Both c and p are fixed and do not depend on the volume of inventory or any sales effort. Notice that without further specification, we have defined a classical newsvendor problem, which will serve as the base model for understanding the trade-offs that can arise in a moral-hazard salesforce incentive context.

The firm and the sales agent follow the typical relationship between a principal and an agent. The firm designs the incentive contract, $t(\cdot)$, for the sales agent and makes the inventory decision, q, based on which the sales agent that sells (or helps sell) the product on behalf of the firm decides the sales effort, e. We assume that neither the sales effort nor the real demand is directly observable. Even though the firm can infer the demand from the sales when there is excess inventory, the real demand would remain unrevealed if it exceeds the inventory level and the excess portion is lost; that is, the information of the real demand is censored by the inventory level. Hence, an incentive contract, $t(\cdot)$, that the firm designs in our model can only be contingent on the realized sales instead of the sales effort or the demand. We assume that the sales agent is risk neutral; however, he has limited wealth and thus can receive only nonnegative payment for any sales realization, that is, $t(\cdot) \geq 0$ is required. Notice that such a specification follows the limited liability assumption that has often been made in the agency literature (see, e.g., Sappington 1983, Park 1995, Kim 1997, Oyer 2000, Chu and Sappington 2009). Exerting a sales effort e is costly for the agent, which is captured by a convex cost function $C(e)$ with $C(0) = 0$ and $C(e) > 0$ for $e > 0$. To ensure an interior solution, we assume $C'(0) = 0$ and $C'(\infty) = \infty$. Finally, without loss of generality, we normalize the agent's reservation utility to zero.

The demand for the product is jointly determined by the agent's sales effort and some uncertain market condition. We use X to represent the stochastic, nonnegative demand and x to represent the realized demand or a real variable. Ex ante, the demand X follows a distribution function $F(x \mid e)$ (with density $f(x \mid e)$) on the interval $[L(e), R(e)]$, where $L(e)$ and $R(e)$ are nondecreasing in e ($R(e)$ can be infinite). We assume that $F(x \mid e)$ satisfies the monotone likelihood ratio property (MLRP) (i.e., $\partial/\partial x(f_e(x \mid e)/f(x \mid e)) \geq 0$, where $f_e(x \mid e)$ is the partial derivative with respect to e); this property implies that the higher the observed demand, the more likely it is due to a large sales effort. (Many common distributions satisfy MLRP. In particular, for the additive ($X = e + \xi$) and multiplicative ($X = e\xi$)

effort cases, $F(x \mid e)$ satisfies MLRP if ξ follows, for instance, the uniform distribution, the normal distribution, the exponential distribution, and most of the beta, gamma, and chi-squared distributions.) Furthermore, we focus only on those demand and effort cost settings where the agent's problem is unimodal. The unimodality condition serves as a regularity condition that will ensure the agent's problem has a unique maximizer and thus enable the classic first-order approach. In our context, this can be *sufficiently* satisfied with those distributions following the convexity distribution function condition (i.e., $F_e(x \mid e) < 0$ and $F_{ee}(x \mid e) \geq 0$) that is commonly assumed in the moral-hazard literature (see, e.g., Holmström 1979, Grossman and Hart 1983, Park 1995, Kim 1997, Jewitt et al. 2008), and it may also be satisfied with other distributions in specific scenarios.

3.2. Formulation

The net utility of the sales agent is the compensation received from the firm minus the cost of effort. Hence, given the incentive contract $t(\cdot)$ and the inventory q, to maximize the expected utility, the agent solves the following optimization problem:

$$\max_{e} \mathbf{E}[t(\min\{q, X\})] - C(e) = \int_{L(e)}^{R(e)} t(\min\{q, x\}) \, \mathrm{d}F(x \mid e) - C(e).$$

Let $e(t, q)$ be the agent's optimal effort decision, and let $u(t, q)$ be the corresponding expected utility. Then, the firm's problem of finding an optimal contract and inventory level can be formulated as

$$\begin{aligned} \max_{t(\cdot), q} \quad & \mathbf{E}[p\min\{q, X\} - t(\min\{q, X\})] - cq \\ & = \int_{L(e)}^{R(e)} \big(p\min\{q, x\} - t(\min\{q, x\})\big) \, \mathrm{d}F(x \mid e) - cq && (1) \\ \text{s.t.} \quad & t(\cdot) \geq 0, && \text{(LL)} \\ & e = e(t, q), && \text{(IC)} \\ & u(t, q) \geq 0. && \text{(IR)} \end{aligned}$$

Equation (1) reflects the firm's desire to maximize the expected difference between the sales profit and the compensation paid to the agent. The first constraint in (1) is the limited liability (LL) constraint, ensuring any possible compensation the agent receives is nonnegative, and the remaining two constraints are the incentive compatibility (IC) and individual rationality (IR) constraints, respectively, indicating that the effort level is optimal for the agent and he is better off participating.

3.3. Specific Demand and Effort Cost

In our analysis, we will first characterize several structural properties with general demand distribution and cost function that satisfy the assumed properties. Then, to better understand the effect of demand censorship, we will focus on two demand–effort relationships: the additive ($X = e + \xi$) and multiplicative ($X = e\xi$) effort cases, where ξ is a random variable representing the uncertain market condition. We further assume that ξ has a uniform distribution $U[a, a + \Delta]$, with $a, \Delta > 0$, and the effort cost follows a quadratic function $C(e) = (1/(2k))e^2$, with $k > 0$. Notice that a is a location factor of the distribution function, whereas Δ not only influences the mean but also reflects the intrinsic variability of the distribution. Such demand and cost settings will enable us to derive closed-form solutions. The additive and multiplicative cases are widely adopted in the salesforce literature as well as in the pricing newsvendor literature (e.g., Rao 1990; Petruzzi and Dada 1999; Agrawal and Seshadri 2000; Chen 2000, 2005; Chen and Xiao 2009). These two cases are specific but can serve as surrogates for understanding the effects of demand censorship on salesforce contracting and inventory decision.

4. Benchmark

In this section, we first explore the first-best solution assuming the sales effort is contractible, and then analyze the case with noncontractible effort but contractible demand.

When the sales effort is contractible, the firm can implement any effort level by directly compensating the agent for the cost. The firm's problem can thus be formulated as

$$\max_{e, q} \mathbf{E}[p\min\{q, X\}] - cq - C(e) = \int_{L(e)}^{R(e)} p\min\{q, x\} \, \mathrm{d}F(x \mid e) - cq - C(e).$$

Taking the derivatives yields the necessary conditions for the first-best solution (we use e^{FB} and q^{FB} to denote the first-best effort and inventory level):

$$\begin{cases} -p\displaystyle\int_{L(e)}^{q} F_e(x \mid e) \, \mathrm{d}x - C'(e) = 0, \\ p(1 - F(q \mid e)) - c = 0. \end{cases} \tag{2}$$

Notice immediately from (2) that the service level under the first-best solution $(e^{\mathrm{FB}}, q^{\mathrm{FB}})$ always equals $(p - c)/p$, and an interior e^{FB} also exists. However, to determine whether (2) has a unique solution is not straightforward. In fact, some structural assumption over the distribution function in the general form would be necessary to guarantee a unique solution. Nevertheless, all our results in the following hold even if (2) admits multiple solutions. In that scenario, the global maximizer will be the first-best solution. Below, we obtain the first-best solution for the additive and multiplicative effort cases as specified in §3 (we use subscripts A and M to indicate the association with these two cases).

Proposition 1. *Suppose* $\xi \sim U[a, a+\Delta]$ *and* $C(e) = (1/(2k))e^2$. (i) *In the additive effort case, the first-best sales effort and inventory level follow:* $e_A^{\mathrm{FB}} = k(p-c)$ *and* $q_A^{\mathrm{FB}} = e_A^{\mathrm{FB}} + a + ((p-c)/p)\Delta$. (ii) *In the multiplicative effort case, the first-best sales effort and inventory level follow:* $e_M^{\mathrm{FB}} = k(p-c)(a+((p-c)/(2p))\Delta)$ *and* $q_M^{\mathrm{FB}} = e_M^{\mathrm{FB}}(a+((p-c)/p)\Delta)$.

When the effort is not contractible, the firm needs to design an incentive contract to motivate the agent. Without demand censorship, the demand will be a natural instrument for contracting. In principle, the firm can design many different types of contracts contingent on the demand, such as the bonus contract, the commission contract, or more sophisticated contracts. Specifically, we focus on the bonus contract that awards the agent a bonus, B, if the realized demand reaches or exceeds some quota, T, and nothing otherwise; that is,

$$t(x) = \begin{cases} B & x \geq T, \\ 0 & x < T. \end{cases}$$

Proposition 2. *Without demand censorship, the bonus contract (weakly) dominates all other types of contracts; the first-best solution can be implemented by an optimized bonus contract* $(B^{\mathrm{FB}}, T^{\mathrm{FB}})$ *if and only if there exists an* $x < R(e^{\mathrm{FB}})$ *such that* $C'(e^{\mathrm{FB}})/C(e^{\mathrm{FB}}) \leq -F_e(x \mid e^{\mathrm{FB}})/(1-F(x \mid e^{\mathrm{FB}}))$.

The bonus contract is effective in our setting. In particular, we show in the proof that to induce any sales effort, there always exists a bonus contract that can yield less (or equal) rent to the agent than other types of contracts. To achieve the first-best solution, however, it is needed that the rate of change in the agent's compensation must not be less than the rate of change in his disutility at the first-best effort level, $C'(e^{\mathrm{FB}})/C(e^{\mathrm{FB}})$. Under a bonus contract, the rate of change in the agent's compensation is $-F_e(T \mid e^{\mathrm{FB}})/(1-F(T \mid e^{\mathrm{FB}}))$, which is increasing in T. Therefore, if the condition in Proposition 2 holds, then one can always find a bonus contract $(B^{\mathrm{FB}}, T^{\mathrm{FB}})$ that induces the first-best effort level. Clearly, if $F(x \mid e)$ follows a uniform distribution (such as the demand distribution in the additive and multiplicative cases as specified in §3), then this condition always holds.

Corollary 1. *Without demand censorship, the first-best solution can always be attained by an optimized bonus contract if* $F(x \mid e)$ *follows a uniform distribution.*

It is important to point out that when the first-best solution can be attained (i.e., the condition in Proposition 2 holds), the bonus contract is often not the unique optimal contract. For instance, it is not difficult to show that for the additive and multiplicative effort cases specified in §3, besides the bonus contracts, there exist commission contracts that can similarly achieve the first-best solutions if the demand information is not censored.

5. The Effects of Demand Censorship

To ensure the implementability of any demand-contingent contract, the availability of accurate demand information is crucial. It is, however, often challenging to reveal the real demand, in particular in environments where insufficient inventory results in lost sales and reliable proxies reflecting the demand are not available. Compared to the demand, sales are relatively easy to verify and thus serve as an important contracting instrument in practice. In this section, we investigate the effects of demand censorship using sales as the contracting instrument.

5.1. The Effect on the Bonus Contract

Under demand censorship, the incentive contract needs to be contingent on the realized sales, $s = \min\{q, x\}$. Specifically, for the bonus contract, now the agent will receive the bonus if and only if the realized sales reach or exceed some sales quota; that is,

$$t(s) = \begin{cases} B & s \geq T, \\ 0 & s < T. \end{cases}$$

Proposition 3. *Under demand censorship, the sales-quota based bonus contract (weakly) dominates all other types of contracts; however, the first-best solution now can be attained only if the condition in Proposition 2 holds and* $T^{\mathrm{FB}} \leq q^{\mathrm{FB}}$*, and in the scenario where the first-best solution is not attainable, the optimal sales quota is equal to the optimal inventory level.*

Proposition 3 first shows that under demand censorship, the bonus contract contingent on the sales data still outperforms all other types of contracts. As a result, we can focus on the bonus contract to reveal the effects of demand censorship. Proposition 3 further reveals that demand censorship will affect the implementability of the first-best solution. It is easy to understand from Proposition 2 that if the first-best inventory level q^{FB} is lower than the quota T^{FB} that is needed to implement the first-best effort, that is, $q^{\mathrm{FB}} < T^{\mathrm{FB}}$, then the bonus contract $(B^{\mathrm{FB}}, T^{\mathrm{FB}})$ would not be implementable under demand censorship because the agent would never receive the bonus. This implies that the first-best solution will not be attainable; moreover, the quota and the inventory level must be adjusted. In fact, we find that in such a scenario, the quota should be set equal to the inventory level at optimum; that is, the agent will receive the bonus only if the demand is strong that depletes the inventory. Such a contract motivates the agent most efficiently, which also synchronizes the firm's salesforce incentive design and inventory decision.

5.2. The Effects on the Optimal Sales Effort and Service Level

From the discussion of the above subsection, demand censorship will obviously affect the optimal sales effort and service level. In particular, we show in the proof of Proposition 3 that the optimal service level under demand censorship is never lower than the first-best service level. However, a thorough assessment of the optimal effort and service level for a general setting is analytically challenging. Thus, we focus on the additive and multiplicative effort cases as specified in §3 to reveal insights about the effects of demand censorship. (A numerical analysis with more complex distribution functions is provided in the online appendix (available at http://dx.doi.org/10.1287/msom.1120.0424) and obtains largely similar results.) Before proceeding, recall from §4 that for these two specific cases, the first-best solution is always attainable if the demand is not censored.

5.2.1. The Additive Effort Case. For the additive case, the sales effort adds to the mean of the demand but has no influence on the variance. We characterize in Proposition 4 the optimal sales effort and inventory decision under demand censorship (as well as the associated sales-quota-based bonus contract).

Proposition 4. *Suppose $X = e + \xi$ with $\xi \sim U[a, a+\Delta]$ and $C(e) = (1/(2k))e^2$. Then, under demand censorship, the optimal sales effort and inventory level follow:*

$$(e_A^*, q_A^*) = \begin{cases} \left(k(p-c),\, e_A^* + a + \dfrac{p-c}{p}\Delta\right) & \text{if } \Delta \le \dfrac{kp(p-c)}{2c}, \\ \left(\dfrac{4p-2c}{p/\Delta + 4/k},\, \dfrac{e_A^*}{2} + a + \Delta\right) & \text{if } \dfrac{kp(p-c)}{2c} < \Delta \le \dfrac{kp^2(p-c)}{c^2}, \\ \left(0,\, a + \dfrac{p-c}{p}\Delta\right) & \text{if } \Delta > \dfrac{kp^2(p-c)}{c^2}, \end{cases} \tag{3}$$

which can be implemented by the following sales-quota-based bonus contract:

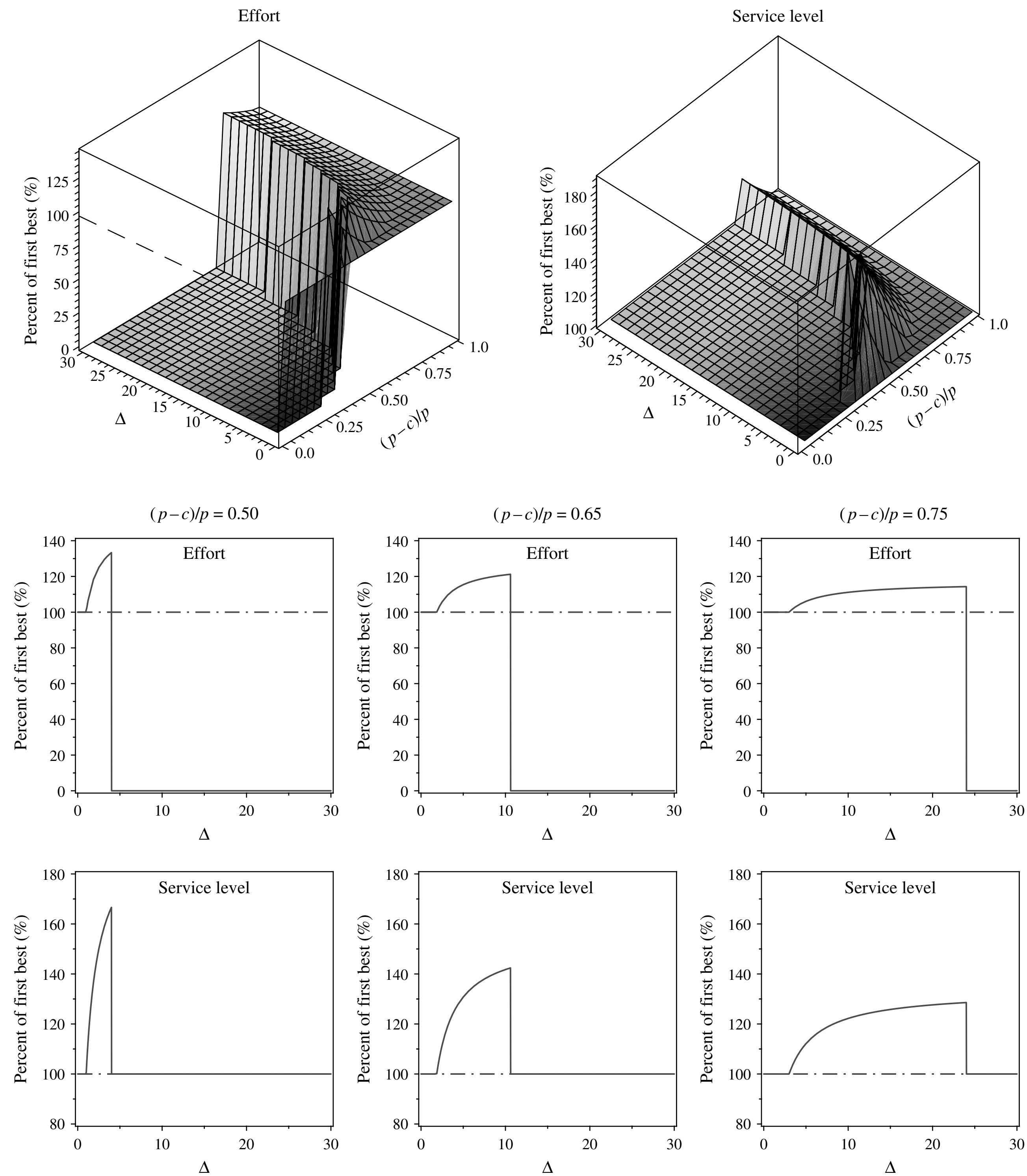
$$(B_A^*, T_A^*) = \begin{cases} \left(\dfrac{\Delta}{k}e_A^*,\, q_A^* + \dfrac{c}{p}\Delta - \dfrac{k(p-c)}{2}\right) & \text{if } \Delta \le \dfrac{kp(p-c)}{2c}, \\ \left(\dfrac{\Delta}{k}e_A^*,\, q_A^*\right) & \text{if } \dfrac{kp(p-c)}{2c} < \Delta \le \dfrac{kp^2(p-c)}{c^2}, \\ (0, N/A) & \text{if } \Delta > \dfrac{kp^2(p-c)}{c^2}. \end{cases} \tag{4}$$

Figure 1 depicts the optimal sales effort and service level based on the results of Proposition 4. Notice that Δ of the distribution function, p and c associated with the sales, and k of the effort cost function all play important roles in determining the sales effort and service level. In particular, the first-best solution is achieved when $\Delta \le (kp(p-c))/(2c)$. When $(kp(p-c))/(2c) < \Delta \le (kp^2(p-c))/c^2$, both the optimal sales effort and service level exceed those under the first-best solution, whereas, in the region with $\Delta > (kp^2(p-c))/c^2$, the optimal sales effort drops abruptly to zero, and the optimal service level reduces to the first-best level. (Note that when $\Delta > (kp^2(p-c))/c^2$, even though the optimal service level coincides with the first-best service level, the absolute volume of demand that is satisfied is obviously lower than that under the first-best solution, given no sales effort is provided.) In the following, we explain the results of Proposition 4 from the perspective of Δ.

We have discussed in §4 that when the demand is contractible, the first-best solution can be achieved through a bonus contract for this additive effort case. Notice that with a large amount of inventory, a bonus contract contingent on sales could function as if it is contingent on the demand. In fact, when the sales quota in the bonus contract that induces the first-best effort is lower than the first-best inventory level, the first-best solution is achievable even in an environment with lost sales. Such a scenario arises when $\Delta \le (kp(p-c))/(2c)$ (see the flat plateau area near the top right corner of the three-dimensional plots as well as the two-dimensional plots in Figure 1).

When Δ increases, for the agent, the marginal return of supplying effort will decrease; and thus, to induce the agent to exert the first-best effort, the firm needs to provide a larger bonus. On the other hand, the firm also wants to extract full rent from the agent, which results in an increase of the sales quota. (A higher sales quota reduces the probability that the bonus is claimed.) In fact, that sales quota increases more than the inventory level (see the first line of (4), where $T_A^* = q_A^* + (c/p)\Delta - (k(p-c))/2$), and it would exceed the inventory level when Δ becomes larger than $(kp(p-c))/(2c)$. In such a scenario, the first-best solution becomes unachievable, and the firm has to align the sales quota with the inventory level. It is interesting to observe that for the additive effort case there is a range of parameters (i.e., $(kp(p-c))/(2c) < \Delta \le (kp^2(p-c))/c^2$) in which it is optimal for the firm to set the sales quota equal to the inventory level and have the induced sales effort and the service level both greater than those under the first-best solution. The firm fully extracts the agency rent. However, as Δ keeps increasing, this distortion will become even larger and more costly

Figure 1 **Illustration for Proposition 4**



Note. $\xi \sim U[a, a+\Delta]$, $C(e) = (1/(2k))e^2$, $p = 2$, $a = 1$, and $k = 1$.

(see the middle peak area of the three-dimensional plots as well as the two-dimensional plots in Figure 1). Recall that the effort cost follows a convex function (for which the firm essentially compensates the agent). Therefore, the benefit the firm obtains from inducing the effort is decreasing as the effort and the inventory level deviate more from the first-best solution.

When Δ exceeds $(kp^2(p-c))c^2$, the firm will find it no longer beneficial to incentivize the agent, and the effort drops abruptly to zero (see the flat valley area near the bottom left corner of the three-dimensional plot as well as the two-dimensional plots in Figure 1). This result is nonintuitive given that for the quadratic cost function, the marginal cost of effort is zero at the zero effort level. Nevertheless, under demand censorship, the firm may need to purposely distort the service level to incentivize the agent and extract full rent through a sales contract, which is associated with a nonnegligible cost. This cost is significant and outweighs the benefit from the sales effort when $\Delta > (kp^2(p-c))/c^2$, and thus the firm will choose to induce zero effort but implement the undistorted, first-best service level. Corollary 2 summarizes these findings (compared to the noncensored demand case).

Corollary 2. *For the additive effort case, when the intrinsic demand variability is relatively small* ($\Delta \le (kp(p-c))/(2c)$), *demand censorship does not affect the optimal effort and service level the firm implements; when the intrinsic demand variability is intermediate* ($(kp(p-c))/(2c) < \Delta \le (kp^2(p-c))/c^2$), *demand censorship increases both the optimal effort and service level; when the intrinsic demand variability is relatively large* ($\Delta > (kp^2(p-c))/c^2$), *demand censorship makes the optimal effort drop to zero but does not affect the optimal service level. Demand censorship does not change the agent's expected utility, which is always zero.*

The above results can be similarly interpreted from the perspective of k, p, or c. Notice that a larger k makes it less costly to exert the sales effort, whereas a larger p or a lower c improves the margin of the sales and thus makes the sales effort more attractive. Hence, if k or p increases or c decreases, the firm will induce a larger effort and, at the same time, set up a larger amount of inventory. The parameter a does not play an important role for the additive effort case, because it is just a location factor of the distribution function. Finally, it is important to note that in contrast to the benchmark with noncensored demand where multiple optimal contracts exist, under demand censorship, the sales-quota-based bonus contract can be uniquely optimal (as far as all relevant compensations are concerned, i.e., those specified in the contract that are receivable by the agent). In particular, for this additive effort case, when $\Delta > (kp(p-c))/(2c)$ (i.e., when the first-best solution is not attainable), the characterized contract in Proposition 4 is uniquely optimal under demand censorship.

5.2.2. The Multiplicative Effort Case. The multiplicative effort case differs from the additive effort case in that the effort influences not only the mean but also the variance of the demand. In Proposition 5, we characterize the optimal solution under demand censorship.

Proposition 5. *Suppose* $X = e\xi$ *with* $\xi \sim U[a, a+\Delta]$ *and* $C(e) = (1/(2k))e^2$. *Then, under demand censorship, the optimal selling effort and inventory level follow:*

$$(e_M^*, q_M^*) = \begin{cases} \left(k(p-c)\left(a+\dfrac{p-c}{2p}\Delta\right), \left(a+\dfrac{p-c}{p}\Delta\right)e_M^*\right) \\ \qquad \text{if } \Delta \le \dfrac{ap}{(3c-p)^+}, \\ \left(\dfrac{kp(8/3(a+\Delta)^2-3a^2)-4kc\Delta(a+\Delta)}{6\Delta}, \dfrac{2}{3}(a+\Delta)e_M^*\right) \\ \qquad \text{if } \dfrac{ap}{(3c-p)^+} < \Delta \le \Delta_M, \\ \left(\dfrac{k\gamma^2(c\Delta-p(a+\Delta-\gamma))}{(a+\Delta)\Delta}, \gamma e_M^*\right) \quad \text{if } \Delta > \Delta_M, \end{cases} \tag{5}$$

which can be implemented by the following sales-quota-based bonus contract:

$$(B_M^*, T_M^*) = \begin{cases} \left(\dfrac{3\Delta(e_M^*)^2}{2k(a+\Delta)}, \dfrac{2(a+\Delta)}{3(a+((p-c)/p)\Delta)}q_M^*\right) \\ \qquad \text{if } \Delta \le \dfrac{ap}{(3c-p)^+}, \\ \left(\dfrac{3\Delta(e_M^*)^2}{2k(a+\Delta)}, q_M^*\right) \\ \qquad \text{if } \dfrac{ap}{(3c-p)^+} < \Delta \le \Delta_M, \\ \left(\dfrac{\Delta(e_M^*)^2}{k\gamma}, q_M^*\right) \quad \text{if } \Delta > \Delta_M, \end{cases} \tag{6}$$

where Δ_M *is the unique solution of*

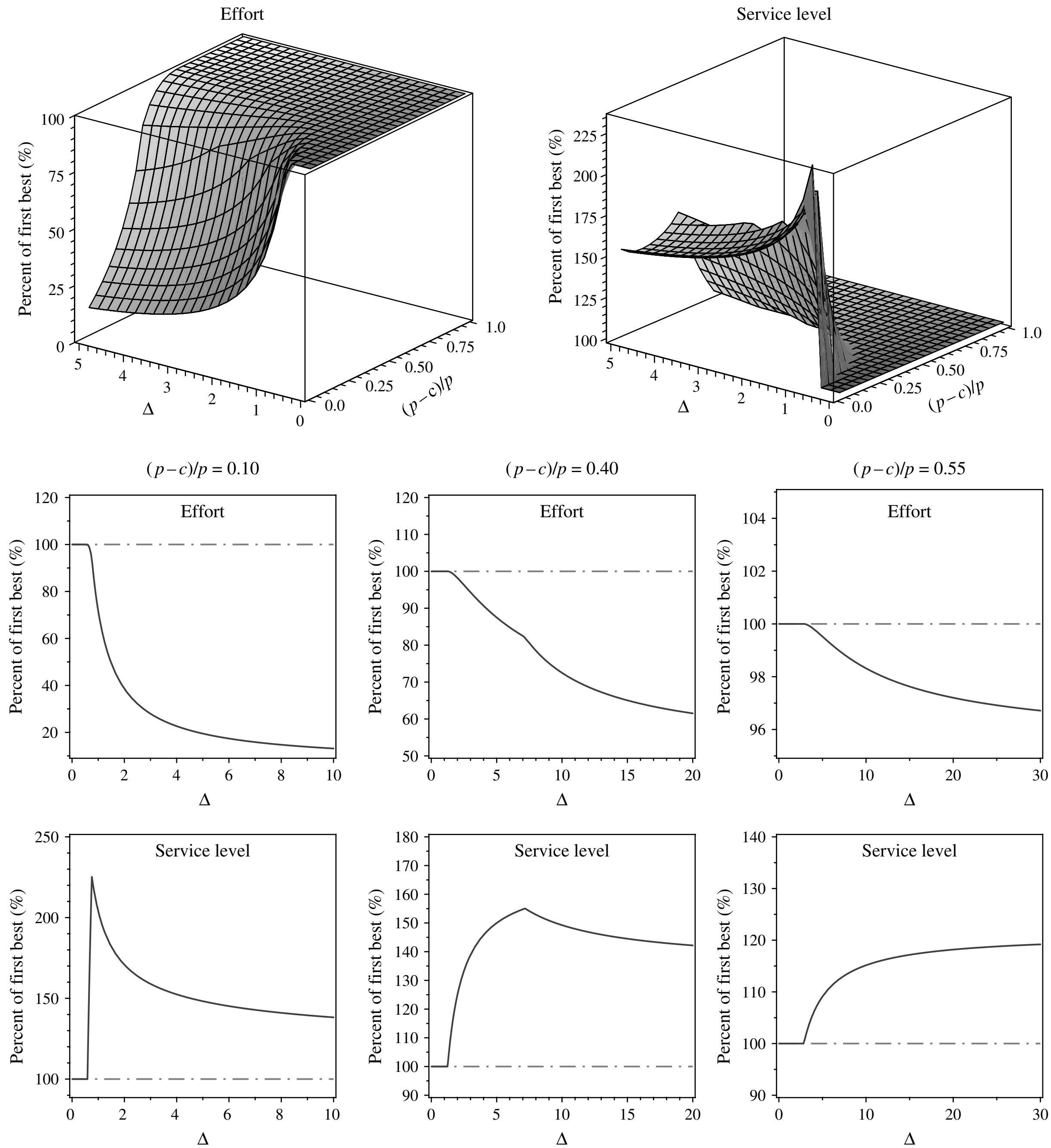
$$\frac{c}{p} = \frac{32(a+\Delta)^2 - 27a^2}{60\Delta(a+\Delta)}$$

on $\mathbb{R}^+$ *if it exists and* ∞ *otherwise, and* γ *is the second largest root of the cubic function:*

$$p\frac{2(a+\Delta)\gamma - \gamma^2 - a^2}{2\Delta} - c\gamma + \frac{2(a+\Delta-\gamma)(p(a+\Delta-\gamma)-c\Delta)\gamma}{(a+\Delta)\Delta} = 0. \tag{7}$$

Comparing Proposition 5 with Proposition 1, we can see that for this multiplicative effort case, when $\Delta \le (ap)/(3c-p)^+$, the first-best solution is achieved. Note that if the profit margin $(p-c)/p$ (which is also the first-best service level) is above 2/3, or identically, if $c \le p/3$, then $\Delta \le (ap)/(3c-p)^+$ always holds regardless of the level of Δ, given that $(ap)/(3c-p)^+$ will go to infinity. Specifically, in such

Figure 2 **Illustration for Proposition 5**



Note. $\xi \sim U[a, a+\Delta]$, $C(e) = (1/(2k))e^2$, $p = 2$, $a = 1$, and $k = 1$.

a case, the quota in the bonus contract that induces the first-best effort is always smaller than the first-best inventory level, and thus demand censorship will not affect the implementation of the first-best solution.

In contrast, if the profit margin is below 2/3 (i.e., $c > p/3$), then under demand censorship, the first-best solution can be achieved only if Δ is small (lower than $(ap)/(3c-p)$) (see the flat areas in Figure 2). In this scenario, the first-best inventory level is still above the quota of the bonus contract that induces the first-best effort. If $\Delta > (ap)/(3c-p)$, then the first-best inventory level will fall below that quota, and as a result, both the bonus contract and the inventory level need to be adjusted in the face of demand censorship. In fact, it will become optimal for the firm to set the quota equal to the inventory level. For the multiplicative effort case, it is interesting to observe that at optimum the firm will always induce a sales effort lower than the first-best effort but set the service level above the first-best level (see the valley areas for the effort and the peak areas for the service level in Figure 2). Furthermore, when Δ is intermediate ($(ap)/(3c-p) < \Delta \le \Delta_M$), the firm overstocks inventory substantially to extract full rent from the agent (the sales quota is set equal to the inventory level; a larger sales quota limits the rent the agent can obtain). However, the cost of inventory distortion will keep increasing as Δ increases. When $\Delta > \Delta_M$, this cost becomes significant, and thus the firm starts to reduce overstocking while surrendering some rent to the agent. The tradeoff between the rent and the inventory distortion is captured by the cubic function, (7). Notice that γ, which is the ratio of q_M^* to e_M^*, has a natural interpretation linked to the service level (i.e., $(\gamma - a)/\Delta$). A larger γ would indicate a larger distortion. For the multiplicative effort case, the firm will always induce some amount of sales effort because the demand would otherwise drop to zero. We summarize these findings in Corollary 3 (compared to the noncensored demand case).

Corollary 3. *For the multiplicative effort case, when the profit margin is relatively large* ($(p-c)/p \ge 2/3$) *or the intrinsic demand variability is relatively small* ($\Delta \le (ap)/(3c-p)$), *demand censorship does not affect the optimal effort and service level the firm implements; in contrast, when the profit margin is relatively small* ($(p-c)/p < 2/3$) *while the intrinsic demand variability is relatively large* ($\Delta > (ap)/(3c-p)$), *demand censorship will lead to a smaller optimal effort but a higher optimal service level. Also, demand censorship can result in a positive expected utility for the agent when the profit margin is relatively small* ($(p-c)/p < 2/3$) *and the intrinsic demand variability is large* ($\Delta > \Delta_M$).

The above results can be similarly interpreted from the perspective of k, p, or c. We can also show that the optimal effort and inventory level increase in k and p and decrease in c. Finally, in contrast to the benchmark with noncensored demand where multiple optimal contracts exist, the characterized contract in Proposition 5 is uniquely optimal under demand censorship when the first-best solution is not attainable (i.e., when $(p-c)/p < 2/3$ and $\Delta > (ap)/(3c-p)$).

5.2.3. Remarks. The above analysis enables us to draw two important observations. First, for both cases, the presence of demand censorship can significantly reduce the region where the first-best solution can be attained, and when the first-best solution is not achievable, the bonus contract can become uniquely optimal. In contrast, multiple optimal contracts exist in the benchmark with noncensored demand, where the first-best solution can always be achieved.

Second and more interestingly, we find that the pattern of the optimal sales effort in the additive effort case has a fundamental difference from that in the multiplicative effort case. In the additive effort case, demand censorship may lead the firm to incentivize the agent to exert an effort greater than the first-best level but may also result in zero effort. Differently, in the multiplicative effort case, the optimal effort level under demand censorship is never higher than the first-best level but always positive.

The key difference between these two cases lies in how the effort influences the demand uncertainty. For the additive effort case, the sales effort increases the mean of the demand but does not affect the variance. Therefore, a larger sales effort will reduce the coefficient of variation that is an important measure of demand uncertainty for inventory decision. Such a result implies that a larger sales effort in general will help reduce the demand–supply mismatch cost. Consequently, when the first-best solution becomes unattainable, it can be more beneficial for the firm to induce more effort than less, as shown in Figure 1. However, to induce a larger effort implies that the firm needs to offer a larger bonus, which risks losing a significant rent to the agent. As a result, in the context with a bonus contract, the firm distorts the service level by overinvesting in inventory significantly to lift the sales quota that triggers the bonus payment. Such an inventory distortion is costly. Beyond certain magnitude, this inventory distortion can drive the firm to induce zero effort, as the demand uncertainty becomes significantly high.

For the multiplicative effort case, the sales effort acts as a scaling factor over both the mean and the variance of the demand. The coefficient of variation is independent of the effort, whereas the variance increases in the effort. Thus, a larger effort in general will make it more challenging to match the supply with the demand. As a result, when the first-best solution becomes unattainable, it is always more beneficial for the firm to induce less effort, as shown in Figure 2.

It is interesting to observe that the effect of demand censorship on the optimal sales effort and service level that the firm implements (relative to the first-best solution that would always be achieved without demand censorship) can differ significantly in these two cases. The additive and multiplicative cases have been widely adopted in the pricing-inventory literature (see, e.g., Petruzzi and Dada 1999, Agrawal and Seshadri 2000). Similar to the price, the sales effort in our study acts as a demand-shaping instrument. Often, for branded products (or major retail chains) that have high consumer recognition, exerting a sales effort to broaden the market awareness of a specific line of products should result in higher customer traffic but may not necessarily increase the uncertainty in the demand. Instead, the uncertainty for sales and inventory planning may mainly come from a forecasting error (for instance, when the classical time series models are used to predict the demand) or some random shock that is independent of the sales effort. The insights we revealed for the additive effort case may thus apply to such situations, where, under demand censorship, a firm could optimally induce a sales effort higher than the first-best level or even a zero-effort depending on the intrinsic demand variability. Differently, for products (or small retailers) with low consumer recognition, a sales effort to expand the market may also scale up the uncertainty in the demand because there may exist significant uncertainty in each customer's purchasing decision. The multiplicative effort case may thus fit better such situations, where, under demand censorship, the optimal sales effort to induce should never exceed the first-best level. An analogous interpretation of the additive and multiplicative cases has been provided by Agrawal and Seshadri (2000) for joint pricing and inventory decisions.

5.3. The Effect on the Value of Contracting with the Sales Agent

In this subsection, we show numerically the effect of demand censorship on the value of contracting with the sales agent. We continue to consider the additive and multiplicative effort cases.

We first compute the firm's maximum profit without contracting with the sales agent, $\underline{\Pi}$. We then compute the firm's profit under the first-best solution, Π^{FB}. The difference between these two profits ($\bar{\Lambda} = \Pi^{\text{FB}} - \underline{\Pi}$) measures the maximum surplus that the firm can gain by contracting with the sales agent. We further compute the firm's profits in three different scenarios under demand censorship. In the first two scenarios, we first solve the bonus contract that implements the first-best solution without demand censorship and the first-best inventory level. If the derived quota is lower than the inventory level, we implement this contract and inventory level directly. In contrast, if the quota is greater than the inventory level, then (i) in the first scenario, we reset the inventory level to the quota and implement the contract and the revised inventory level; (ii) in the second scenario, we reset the quota to the first-best inventory level and implement the revised contract and the first-best inventory level. The first scenario corresponds to a situation where a firm's salesforce contracting decision does not consider demand censorship and it is not coordinated with the inventory decision; the latter is a passive decision that is made contingent on the former. In contrast, the second scenario corresponds to a situation where a firm's inventory decision does not consider the effect of demand censorship on salesforce contracting and the latter decision is a passive one that is contingent on the inventory decision. In the last scenario, we solve the optimal bonus contract and the inventory decision simultaneously (following §5.2). In all of these scenarios, we assume that the sales agent always takes demand censorship into account when deciding the effort level. We compute the firm's profits Π^{NCI}, Π^{NCII}, and Π^{C}, corresponding to the above three scenarios in sequence, and the differences $\Lambda^{NCI} = \Pi^{NCI} - \underline{\Pi}$, $\Lambda^{NCII} = \Pi^{NCII} - \underline{\Pi}$, and $\Lambda^{C} = \Pi^{C} - \underline{\Pi}$, respectively.

Table 1 reports the results of our numerical study. In the experiments, we set the location parameter a of the uniform distribution at 1 but vary Δ from 1 to 5. Notice that with this setting, the coefficient of variation of the uniform distribution ranges from 0.19 (when $\Delta = 1$) to 0.41 (when $\Delta = 5$); thus, the intrinsic demand variability in this set of experiments is not large. We also vary the profit margin $(p - c)/p$ from 10% to 85%.

From Table 1, we can observe that the effect of demand censorship can be very significant when the profit margin is small. For instance, for the additive effort case, when the profit margin is below 25%, Λ^{NCI} and Λ^{NCII} are all negative, and Λ^{C} is zero for any Δ we choose; that is, contracting with the sales agent does not contribute any (positive) value to the firm under demand censorship. In particular, notice that if the firm does not correctly take demand censorship into account, the value of contracting with the sales agent can be significantly negative (Λ^{NCI} can be up to -3.85). In contrast, $\bar{\Lambda}$, the value of contracting with the sales agent under the first-best solution (or identically, the value of contracting with the sales agent if demand is not censored), is always positive and represents an 8% to 22% increment of the firm's profit (i.e., $((\Pi^{\text{FB}} - \underline{\Pi})/\underline{\Pi}) \times 100\%$) when the profit margin is below 25% in our experiments. Intuitively, for a small profit margin, the inventory that the firm installs is generally low relative to the quota that is needed to appropriately contract with the agent, and thus lost sales can occur often. Furthermore, a small profit margin implies a large c relative to p. As a result, a misaligned inventory investment can cause significant losses. From Table 1, we can observe that for the additive effort case, as the profit margin increases, the negative effect of demand censorship weakens; however, it can still erode a substantial proportion of the agent's value or lead to a negative value when Δ is relatively large (unless the profit margin reaches 85% in our experiments). Similar observations can be drawn from the data of the multiplicative effort case. Note that the gross margins of most manufacturing firms as well as retail chains are typically below 50% in practice. Our results indicate that demand censorship can cause a significant impact on firms, especially those with relatively small gross margins, and it is important to correctly take demand censorship into account in their sales and inventory planning.

Table 1 Comparison of the Values from Contracting with the Sales Agent, Captured by $\begin{pmatrix} \Lambda^{NCI} & \Lambda^{NCII} \\ \Lambda^{C} & \bar{\Lambda} \end{pmatrix}$

$(p-c)/p$	$\Delta=1$	$\Delta=2$	$\Delta=3$	$\Delta=4$	$\Delta=5$
			Additive effort case		
0.10	$\begin{pmatrix} -0.62 & -0.14 \\ 0.00 & 0.02 \end{pmatrix}$	$\begin{pmatrix} -1.43 & -0.32 \\ 0.00 & 0.02 \end{pmatrix}$	$\begin{pmatrix} -2.23 & -0.50 \\ 0.00 & 0.02 \end{pmatrix}$	$\begin{pmatrix} -3.04 & -0.68 \\ 0.00 & 0.02 \end{pmatrix}$	$\begin{pmatrix} -3.85 & -0.86 \\ 0.00 & 0.02 \end{pmatrix}$
0.25	$\begin{pmatrix} -0.13 & -0.13 \\ 0.00 & 0.13 \end{pmatrix}$	$\begin{pmatrix} -0.66 & -0.50 \\ 0.00 & 0.13 \end{pmatrix}$	$\begin{pmatrix} -1.21 & -0.88 \\ 0.00 & 0.13 \end{pmatrix}$	$\begin{pmatrix} -1.77 & -1.25 \\ 0.00 & 0.13 \end{pmatrix}$	$\begin{pmatrix} -2.33 & -1.63 \\ 0.00 & 0.13 \end{pmatrix}$
0.40	$\begin{pmatrix} 0.28 & 0.16 \\ 0.29 & 0.32 \end{pmatrix}$	$\begin{pmatrix} 0.00 & -0.32 \\ 0.06 & 0.32 \end{pmatrix}$	$\begin{pmatrix} -0.33 & -0.80 \\ 0.00 & 0.32 \end{pmatrix}$	$\begin{pmatrix} -0.68 & -1.28 \\ 0.00 & 0.32 \end{pmatrix}$	$\begin{pmatrix} -1.03 & -1.76 \\ 0.00 & 0.32 \end{pmatrix}$
0.55	$\begin{pmatrix} 0.61 & 0.61 \\ 0.61 & 0.61 \end{pmatrix}$	$\begin{pmatrix} 0.54 & 0.22 \\ 0.56 & 0.61 \end{pmatrix}$	$\begin{pmatrix} 0.39 & -0.28 \\ 0.42 & 0.61 \end{pmatrix}$	$\begin{pmatrix} 0.21 & -0.77 \\ 0.26 & 0.61 \end{pmatrix}$	$\begin{pmatrix} 0.03 & -1.27 \\ 0.08 & 0.61 \end{pmatrix}$
0.70	$\begin{pmatrix} 0.98 & 0.98 \\ 0.98 & 0.98 \end{pmatrix}$	$\begin{pmatrix} 0.98 & 0.98 \\ 0.98 & 0.98 \end{pmatrix}$	$\begin{pmatrix} 0.97 & 0.70 \\ 0.97 & 0.98 \end{pmatrix}$	$\begin{pmatrix} 0.92 & 0.28 \\ 0.92 & 0.98 \end{pmatrix}$	$\begin{pmatrix} 0.85 & -0.14 \\ 0.86 & 0.98 \end{pmatrix}$
0.85	$\begin{pmatrix} 1.45 & 1.45 \\ 1.45 & 1.45 \end{pmatrix}$	$\begin{pmatrix} 1.45 & 1.45 \\ 1.45 & 1.45 \end{pmatrix}$	$\begin{pmatrix} 1.45 & 1.45 \\ 1.45 & 1.45 \end{pmatrix}$	$\begin{pmatrix} 1.45 & 1.45 \\ 1.45 & 1.45 \end{pmatrix}$	$\begin{pmatrix} 1.45 & 1.45 \\ 1.45 & 1.45 \end{pmatrix}$
			Multiplicative effort case		
0.10	$\begin{pmatrix} 0.01 & 0.01 \\ 0.02 & 0.02 \end{pmatrix}$	$\begin{pmatrix} -0.05 & -0.02 \\ 0.01 & 0.02 \end{pmatrix}$	$\begin{pmatrix} -0.12 & -0.06 \\ 0.01 & 0.03 \end{pmatrix}$	$\begin{pmatrix} -0.20 & -0.09 \\ 0.01 & 0.03 \end{pmatrix}$	$\begin{pmatrix} -0.28 & -0.13 \\ 0.01 & 0.03 \end{pmatrix}$
0.25	$\begin{pmatrix} 0.15 & 0.13 \\ 0.15 & 0.16 \end{pmatrix}$	$\begin{pmatrix} 0.12 & -0.00 \\ 0.13 & 0.20 \end{pmatrix}$	$\begin{pmatrix} 0.04 & -0.14 \\ 0.12 & 0.24 \end{pmatrix}$	$\begin{pmatrix} -0.05 & -0.28 \\ 0.12 & 0.28 \end{pmatrix}$	$\begin{pmatrix} -0.17 & -0.44 \\ 0.12 & 0.33 \end{pmatrix}$
0.40	$\begin{pmatrix} 0.46 & 0.46 \\ 0.46 & 0.46 \end{pmatrix}$	$\begin{pmatrix} 0.60 & 0.42 \\ 0.61 & 0.63 \end{pmatrix}$	$\begin{pmatrix} 0.73 & 0.30 \\ 0.73 & 0.82 \end{pmatrix}$	$\begin{pmatrix} 0.84 & 0.16 \\ 0.85 & 1.04 \end{pmatrix}$	$\begin{pmatrix} 0.96 & 0.00 \\ 0.98 & 1.28 \end{pmatrix}$
0.55	$\begin{pmatrix} 0.98 & 0.98 \\ 0.98 & 0.98 \end{pmatrix}$	$\begin{pmatrix} 1.45 & 1.45 \\ 1.45 & 1.45 \end{pmatrix}$	$\begin{pmatrix} 2.01 & 1.98 \\ 2.01 & 2.02 \end{pmatrix}$	$\begin{pmatrix} 2.66 & 2.33 \\ 2.66 & 2.67 \end{pmatrix}$	$\begin{pmatrix} 3.38 & 2.73 \\ 3.38 & 3.41 \end{pmatrix}$
0.70	$\begin{pmatrix} 1.79 & 1.79 \\ 1.79 & 1.79 \end{pmatrix}$	$\begin{pmatrix} 2.83 & 2.83 \\ 2.83 & 2.83 \end{pmatrix}$	$\begin{pmatrix} 4.12 & 4.12 \\ 4.12 & 4.12 \end{pmatrix}$	$\begin{pmatrix} 5.64 & 5.64 \\ 5.64 & 5.64 \end{pmatrix}$	$\begin{pmatrix} 7.41 & 7.41 \\ 7.41 & 7.41 \end{pmatrix}$
0.85	$\begin{pmatrix} 2.93 & 2.93 \\ 2.93 & 2.93 \end{pmatrix}$	$\begin{pmatrix} 4.95 & 4.95 \\ 4.95 & 4.95 \end{pmatrix}$	$\begin{pmatrix} 7.48 & 7.48 \\ 7.48 & 7.48 \end{pmatrix}$	$\begin{pmatrix} 10.5 & 10.5 \\ 10.5 & 10.5 \end{pmatrix}$	$\begin{pmatrix} 14.1 & 14.1 \\ 14.1 & 14.1 \end{pmatrix}$

Note. The common setting follows: $\xi \sim U[a, a+\Delta]$, $C(e) = (1/(2k))e^2$, $p = 2$, $a = 1$, and $k = 1$.

6. Extensions and Discussions

In this section, we discuss three possible extensions of our model, including adverse selection, different risk preferences, and multiple sales agents.

6.1. Adverse Selection

For the analysis in the above sections, we have assumed that the firm and the sales agent have the same information set with respect to the market condition; that is, we have explored a pure moral-hazard problem. In this subsection, we discuss a setting where the sales agent has some private information of the market condition. Note that such a combined adverse-selection and moral-hazard problem is technically challenging. For tractability, we consider the additive effort case as specified in §3, that is, $X = e + \xi$ with $\xi \sim U[a, a+\Delta]$ and $C(e) = (1/2k)e^2$ (the multiplicative effort case is much more difficult to analyze). To incorporate the notion that the sales agent knows more about the market, we assume that the location parameter of the uniform distribution can be either $a = a_H$ with probability ρ or $a = a_L$ with probability $1-\rho$; the sales agent knows the true market condition, whereas the firm knows only the prior distribution. Now, the firm needs to design a menu of two incentive contracts targeting the corresponding market scenarios. We formulate this adverse-selection and moral-hazard problem in the online appendix. Recall from §4 that $e^{\mathrm{FB}} = k(p-c)$ for the additive effort case. The following proposition shows that when the difference between the two market scenarios is large, the bonus contract continues to be optimal even if the agent has private information.

Proposition 6. *When $a_H - a_L \geq 2e^{\mathrm{FB}} + \Delta$, there exists an optimal menu of sales-quota-based bonus contracts that yield the firm the largest expected profit under demand censorship.*

We explain the intuition of Proposition 6 as follows. Without information asymmetry, the bonus contract can yield the agent the least rent to implement the optimal effort and inventory combination in each market scenario. However, if the sales agent privately observes the true market scenario while the firm knows only the prior distribution, then the agent may have an incentive to always report a weak market scenario. Thus, the firm needs to take into account the agent's incentive when designing the salesforce contracts. By the classical mechanism design theory, under the optimal menu of contracts, the system should achieve efficiency (i.e., without any information caused distortion) when the agent observes a strong market scenario even though he may obtain an information rent. Given that the bonus contract yields the agent the least moral-hazard rent to implement any effort, the bonus contract must be optimal for the strong market scenario. The optimal contract that targets the weak market scenario, however, may not necessarily be a bonus contract. The firm faces a trade-off of distorting the contract targeting the weak market scenario, which loses efficiency, versus surrendering a large information rent to the agent under the strong market scenario. Nevertheless, we find that as long as the difference between the two market conditions is sufficiently large (i.e., $a_H - a_L \geq 2e^{\mathrm{FB}} + \Delta$), the information rent surrendered to the agent under the strong market scenario is determined by the largest possible payment to the agent under the contract that targets the weak market scenario. Note that among all the contracts that induce the same effort level and under which the largest possible payment to the agent is the same, the bonus contract will yield the least moral-hazard rent to the agent. As a result, when $a_H - a_L \geq 2e^{\mathrm{FB}} + \Delta$, the bonus contract is also optimal for the weak market scenario.

Clearly, because the bonus contract continues to be optimal, the observations we obtained in §5.2.1 will still arise in this additive effort case with asymmetric information; that is, demand censorship can lead to a larger optimal effort as well as a higher service level, and scenarios also exist where demand censorship can result in zero effort. Note, however, that with asymmetric information, demand censorship may introduce more peculiar effects. When the difference between the two market conditions $(a_H - a_L)$ is small, the information rent surrendered to the agent under the strong market scenario depends not just on the largest possible payment to the agent under the weak market scenario; instead, it depends on the payment schedule of all possible sales realizations. As a result, the bonus contract may no longer be optimal for the weak market scenario. Solving the optimal menu of contracts in such a situation, in fact, is significantly challenging, which warrants future exploration.

6.2. Generalized Limited Liability and Risk Aversion

We have assumed that the sales agent is risk neutral but has limited liability. Note that contracts under which one or more parties have limited liability are very common in practice for both individuals and companies (Sappington 1983). In our study, specifically, the lowest possible payment that the sales agent receives is zero. One natural extension would be to assume a minimum payment t_0 (> 0) that the agent needs to receive. Notice that to incorporate this extension, we can add a fixed transfer payment t_0 to the bonus contract; that is, the agent receives t_0 if the sales are lower than the quota T and $t_0 + B$ otherwise. Such a change would not affect the analysis because the fixed transfer payment does not change the incentive structure under a bonus contract. However, if the agent has some wealth and can be punished for poor sales realizations (i.e., if the agent needs to compensate the firm for low sales realizations), then the first-best solution will be easier to induce, or, identically, the effect of demand censorship will be alleviated. In the extreme, if the agent has unlimited wealth, then the first-best solution can always be achieved by selling the firm to the agent.

One might also assume that the sales agent is risk averse. It is well known from the literature that when the agent is risk averse, the optimal contract will likely be less steeply sloped because steeply sloped contracts (such as the bonus contract) can impose considerable risk on the agent. In that scenario, it might become more beneficial for the firm to award the agent a commission contingent on the sales realization in addition to a bonus at the inventory boundary. Nevertheless, the insights revealed in our study with respect to demand censorship will likely hold for the setting with a risk-averse sales agent. Demand censorship will cause frictions in the salesforce incentive design as well as distortions in the inventory decision, and the effects can differ for different effort–demand relationships. Formally analyzing the problem with a risk-averse sales agent and demand censorship is, however, substantially involved.

6.3. Multiple Sales Agents

So far, we have assumed that there is only one sales agent and shown that demand censorship can cause a friction in the salesforce incentive design as the firm is not able to observe the full spectrum of demand realizations. Note that the agent in our study should not be treated literally as an individual salesperson; instead, because we focus on how a firm can coordinate the sales planning and inventory decisions, the sales agent in our study shall represent a sales group or a third-party sales agency. In practice, however, it is still possible that a firm may contract with multiple sales agents who independently generate sales, and the firm may install a common inventory pool that the agents share to fulfill their demands. Note that in such a situation, first, a pooling effect may arise due to inventory sharing, and the demand generated by each individual agent might be less censored. Second, the firm can potentially design the incentive contracts for the agents depending on each other's performance, which thus introduces more instruments for contracting. In the online appendix, we provide a specific example where the firm contracts with two independent sales agents who share a common inventory pool; we show that the effect of demand censorship can be alleviated by designing the incentive contracts depending on each agent's performance. A complete investigation with multiple sales agents under a general setting is beyond the scope of our study, which is interesting for future research.

7. Conclusion

In this paper, we investigate the optimal salesforce contract and inventory decision under demand censorship. We reveal that demand censorship can affect the implementation of the first-best solution because the firm may not be able to ideally incentivize the sales agent given that the information of large demand realizations will be censored by the inventory level. We show that in such an environment, the sales-quota-based bonus contract can perform more effectively than other types of contracts, and when the first-best solution is not attainable, the sales quota should be set equal to the inventory level at optimum.

Furthermore, by focusing on the additive and multiplicative effort cases, we find that demand censorship can have a significant negative impact on a firm's performance, and it is especially important for those firms in practice (either retailers or manufacturers) that have a relatively small profit margin and large demand uncertainty to take demand censorship into account. Demand censorship can also have structurally different impacts on the optimal salesforce incentive design and inventory decision in the additive and multiplicative effort cases. In particular, for the additive effort case, which corresponds to business environments where the sales effort can boost the demand without making it more variable, we find that a firm under demand censorship should sometimes induce a sales effort and maintain a service level both greater than those under the first-best solution, and scenarios also exist where zero effort should be induced. In contrast, for the multiplicative effort case, which corresponds to business environments where the sales effort can boost the demand as well as its variability, we find that a firm often should induce a smaller sales effort but maintain a larger service level than those under the first-best solution. These insights can be useful for firms to properly design their salesforce contracts and determine their inventory levels.

Electronic Companion

An electronic companion to this paper is available as part of the online version at http://dx.doi.org/10.1287/msom.1120.0424.

Acknowledgments

The authors thank editor Steve Graves, the associate editor, and the two referees for constructive comments that greatly improved the paper. They also thank Fangruo Chen, Stephen Gilbert, and the seminar participants at the 2011 INFORMS Annual Conference, 2012 MSOM Society Annual Conference, and 2012 Chinese Scholars Association in Management Science and Engineering Annual Conference for helpful discussion and suggestions.

References

Agrawal V, Seshadri S (2000) Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem. *Manufacturing Service Oper. Management* 2(4):410–423.

Albers S, Mantrala M (2008) Models for sales management decisions. Wierenga B, ed. *Handbook of Marketing Decision Models* (Springer, Berlin), 163–210.

Anderson ET, Fitzsimons GJ, Simester D (2006) Measuring and mitigating the costs of stockouts. *Management Sci.* 52(11):1751–1763.

Basu A, Lal R, Srinivasan V, Staelin R (1985) Salesforce compensation plans: An agency theoretic perspective. *Marketing Sci.* 4(4):267–291.

Bensoussan A, Çakanyıldırım M, Sethi SP (2007) A multiperiod newsvendor problem with partially observed demand. *Math. Oper. Res.* 32(2):322–344.

Chen F (2000) Sales-force incentives and inventory management. *Manufacturing Service Oper. Management* 2(2):186–202.

Chen F (2005) Salesforce incentives, market information, and production/inventory planning. *Management Sci.* 51(1):60–75.

Chen F, Xiao W (2009) On compensation schemes for sales agents with superior market information. Working paper, New York University, New York.

Chen L, Plambeck EL (2008) Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing Service Oper. Management.* 10(2):236–256.

Chen Y, Xiao W (2012) Impact of reseller's forecasting accuracy on channel member performance. *Production Oper. Management* 21(6):1075–1089.

Chu LY, Sappington DE (2009) Implementing high-powered contracts to motivate intertemporal effort supply. *RAND J. Econom.* 40(2):296–316.

Ding X, Puterman ML, Bisi A (2002) The censored newsvendor and the optimal acquisition of information. *Oper. Res.* 50(3):517–527.

Emmelhainz MA, Stock RJ, Emmelhainz LW (1991) Consumer responses to retail stock-outs. *J. Retailing* 67(2):138–147.

Grossman SJ, Hart O (1983) An analysis of the principal-agent problem. *Econometrica* 51(1):7–45.

Gruen TW, Corsten DS, Bharadwaj S (2002) Retail out-of-stocks: A worldwide examination of extent, causes and consumer responses. *The Food Institute Forum*.

Heese HS, Swaminathan JM (2010) Inventory and sales effort management under unobservable lost sales. *Eur. J. Oper. Res.* 207(3):1263–1268.

Holmström B (1979) Moral hazard and observability. *Bell J. Econom.* 10(1):74–91.

Hopp W, Iravani SMR, Liu Z (2010) The role of wholesale-salespersons and incentive plans in promoting supply chain performance. Working paper, Northwestern University, Evanston, IL.

Innes RD (1990) Limited liability and incentives contracting with ex-ante action choices. *J. Econom. Theory* 52(1):45–67.

Jerath K, Netessine S, Zhang ZJ (2010) Can we all get along? Incentive contracts to bridge the marketing and operations divide. Working paper, Carnegie Mellon University, Pittsburgh.

Jewitt I, Kadan O, Swinkels JM (2008) Moral hazard with bounded payments. *J. Econom. Theory* 143(1):59–82.

Johnson S (2010) Product management and marketing survey. *The Pragmatic Marketer* 8(1):7–14.

Joseph K, Kalwani MU (1998) The role of bonus pay in salesforce compensation plans. *Indust. Marketing Management* 27(2): 147–159.

Khanjari N, Iravani S, Shin H (2012) The impact of the manufacturer-hired sales agent on a supply chain with information asymmetry. Working paper, Northwestern University, Evanston, IL.

Kim SK (1997) Limited liability and bonus contracts. *J. Econom. Management Strategy* 6(4):899–913.

Lal R, Srinivasan V (1993) Compensation plans for single- and multi-product salesforces: An application of the Holmström-Milgrom model. *Management Sci.* 39(7):777–793.

Lal R, Staelin R (1986) Salesforce compensation plans in environments with asymmetric information. *Marketing Sci.* 5(3): 179–198.

Lariviere MA, Porteus E (1999) Stalking information: Bayesian inventory management with unobserved lost sales. *Management Sci.* 45(3):346–363.

Lu X, Song JS, Zhu K (2006) Analysis of perishable inventory systems with censored data. *Oper. Res.* 56(4): 1034–1038.

Oyer P (2000) A theory of sales quotas with limited liability and rent sharing. *J. Labor Econom.* 18(3):405–426.

Park ES (1995) Incentive contracting under limited liability. *J. Econom. Management Strategy* 4(3):477–490.

Petruzzi NC, Dada M (1999) Pricing and the newsvendor problem: A review with extensions. *Oper. Res.* 47(2):183–194.

Poblete J, Spulber D (2012) The form of incentive contracts: Agency with moral hazard, risk neutrality and limited liability. *RAND J. Econom.* 43(2):215–234.

Raju JS, Srinivasan V (1996) Quota-based compensation plans for multiterritory heterogeneous salesforces. *Management Sci.* 42(10):1454–1462.

Rao R (1990) Compensating heterogeneous salesforces: Some explicit solutions. *Marketing Sci.* 9(4):319–341.

Sappington D (1983) Limited liability contracts between principal and agent. *J. Econom. Theory* 29(1):1–21.

Schary PB, Christopher M (1979) The anatomy of a stock-out. *J. Retailing* 55(2):59–70.

Zoltners A, Sinha P, Lorimer SE (2008) Sales force effectiveness: A framework for researchers and practitioners. *J. Personal Selling Sales Management* 28(2):115–131.