Incentive Design for Operations-Marketing Multitasking

Tinglong Dai[*]   Rongzhu Ke[†]   Christopher Thomas Ryan[‡]

[*]Carey Business School, Johns Hopkins University, Baltimore, Maryland 21202, dai@jhu.edu
[†]Department of Economics, Hong Kong Baptist University, Kowloon, Hong Kong, rongzhuke@hkbu.edu.hk
[‡]Booth School of Business, University of Chicago, Chicago, Illinois 60637, chris.ryan@chicagobooth.edu

A firm hires an agent (e.g., store manager) to undertake both operational and marketing activities for a product. Marketing activities boost demand, but for demand to translate into sales, operational effort is required to maintain adequate inventory. The firm designs a compensation plan to induce the agent to put appropriate effort into both marketing and operations, with the additional challenge of "demand censoring" (i.e., demand in excess of available inventory is unobservable). We formulate this incentive-design problem using a moral hazard principal-agent framework with a multitasking agent subject to a censored signal. We develop a bang-bang approach, with a general optimality structure applicable to a broad class of incentive-design problems. Using this approach, we characterize the optimal compensation plan, with a bonus region resembling a "mast" and "sail," such that a bonus is paid when either all inventory above a threshold is sold or the sales quantity meets an inventory-dependent target. This structure implies nonmonotonicity where the agent can be less likely to receive a bonus for achieving a better outcome. Furthermore, we find the relationship between inventory and demand outcomes can be either complementary or substitutive in the optimal compensation plan. With practical implementability in mind, we approximate the optimal mast-and-sail contract by seeking simple contracts resembling it. We find a type of monotone contract without a "mast" can perform well, but only if the "sail" properly reflects the inventory-sales relationship. By contrast, simple contracts with both a "mast" and a linearized "sail" consistently achieve near-optimal performance.

Key words: Marketing-operations interface, multitasking, retail operations, principal-agent problems.
History: This version: March 11, 2019.

1. Introduction

Multitasking between operations and marketing characterizes many of the job functions in the world of retailing, particularly among store managers. A search on the online employment platform Monster.com returns over 8,500 retail store manager job listings requiring "multitasking" as a core skill. According to DeHoratius and Raman (2007, p. 523), the store manager is a "*multitasking agent* who allocates effort to different activities based on the rewards that accrue from, and the cost of pursuing, each of these activities" (emphasis added). This paper focuses on two activities of a store manager: (a) bolstering customer demand and (b) ensuring inventory can be put in the hands of customers, where it belongs, instead of being misplaced, damaged, spoiled, or stolen through

mismanagement.[1] Balancing these two distinct families of tasks—increasing demand through marketing efforts and maintaining customer access to inventory through operational efforts—challenges store managers to allocate their time and energy effectively.

Outside of the retail sector, numerous settings involve multitasking between operations and marketing. In the global health setting, for example, private agencies engage in delivering and administering vaccines to children in some of the hardest-to-reach places in the world. The success of their work depends not only on effective campaigns to raise public awareness (marketing), but also on managing a "cold" chain system from freezer to freezer to ensure the safety and efficacy of vaccines and putting these vaccines in the hands of those who need it (operations). The 2012 BBC documentary *Ewan McGregor: Cold Chain Mission* depicts this latter aspect.

Drawing from the above motivating scenarios, we study a firm's incentive-design problem when employing a multitasking agent with two concurrent job functions: operations and marketing. Marketing effort increases demand for the product. Operational effort makes the most of the store's resources, particularly in putting inventory in the hands of customers that demand it. Of course, operations and marketing are inextricably tied together. Indeed, to translate demand into sales, demand must be accompanied by a sufficiently abundant available inventory level that is influenced by operational effort. Likewise, an amply available inventory level may be constrained by weak demand due to poor marketing. The outcome of matching demand with available inventory (i.e., sales) depends on the weaker link of this pair. When available inventory is the weaker link, unmet demand is lost and unobservable, a phenomenon known as *demand censoring*. Demand censoring is widely observed in practice and well studied in economics (e.g., Conlon and Mortimer 2013), marketing (e.g., Anupindi et al. 1998), and operations management (e.g., Besbes and Muharremoglu 2012). In this operations-marketing multitasking setting, the goal in designing a compensation plan is to induce the store manager to strike the right balance (from the perspective of the principal) between operational and marketing effort in light of demand censoring.

The empirical operations management literature (e.g., DeHoratius and Raman 2007) has recognized the importance of the operations-marketing multitasking problem. These studies detail how practitioners have outlined an incentive-design approach to address the issue: provide a bonus for a high realization of sales and a high realization of available inventory (or, equivalently, a small realization of loss in inventory due to mismanagement). To the best of our knowledge, these

---

[1] Among the main sources of the discrepancy of recorded inventory and available inventory are shrinkage and misplacement (Atalı et al. 2009). Beck and Peacock (2009) estimate that retailers around the globe suffer a $232 billion annual loss from inventory shrinkage, which is approximately equal to the GDP of Finland. Ton and Raman (2004) find that one out of six customers seeking help from store associates failed to find products because they were misplaced, a phenomenon referred to as "phantom stockouts."

incentive-design problems have not been formally analyzed using a rigorous moral-hazard principal-agent framework. Undertaking this analysis is the focus of this paper. In our model, the agent has two types of effort unobservable to the principal: (1) marketing effort, which increases demand; and (2) operational effort, which increases available inventory. Both effort types affect their associated outcomes *stochastically* so that the firm cannot directly infer agent efforts by observing demand and inventory outcomes. Demand censoring further cripples the principal's observability of agent efforts. Due to censoring, the operational and marketing sides affect the firm's observability differently: whereas unsold inventory can be observed and used as the basis for compensation, unmet demand cannot be observed. Indeed, the two types of efforts may yield only one observable signal when realized demand exceeds realized inventory.

We refer to the problem at hand as a *multitasking store manager problem*, while noting that it applies to other scenarios involving multitasking between operational and marketing activities. For instance, operational effort may be tied to other forms of available capacity mobilized by agent effort to meet demand. The store manager problem is challenging to solve, in part because the validity of the first-order approach, the standard procedure used in deriving the optimal compensation plan for a moral hazard problem, is questionable. Laffont and Martimort (2009, p. 200) point out that "the first-order approach has been one of the most debated issues in contract theory" and "when the first-order approach is not valid, using it can be very misleading." In particular, the convex distribution function condition (CDFC), often assumed in the moral hazard principal-agent literature to support the first-order approach, is satisfied by essentially *none* of the familiar stochastic distributions. The validity of the first-order approach is particularly troubling under a multitasking setting, with a multidimensional effort and a multidimensional output signal.[2] The presence of demand censoring in our setting weakens the observability of the output signal and makes characterizing the optimal compensation plan even more vexing.

To overcome these technical challenges, we develop a "bang-bang" approach that applies to a broad class of incentive-design problems with finitely many actions and a risk-neutral principal and agent. Through a reformulation of the problem, our approach reduces the contracting problem to a finite-dimensional optimization problem, averting the need to apply the first-order approach. It also significantly relaxes conditions needed to establish optimality, allowing for most of the commonly used families of distributions on both the operational and marketing sides. Using this approach, we characterize the optimal compensation plan for a multitasking agent subject to a censored signal. In essence, the contract weighs the informativeness of each combination of inventory and

---

[2] As a subset of the moral hazard principal-agent literature, the multitasking literature (e.g., Holmstrom and Milgrom 1991; Feltham and Xie 1994; Dewatripont et al. 1999; Datar et al. 2001) mostly focuses on deriving optimal parameters of linear compensation schemes, without establishing the optimality of such linear schemes.

sales outcomes and specifies a threshold above which the store manager receives a bonus. Such a threshold corresponds to a level at which the informativeness of the realized outcome *triggers* the belief that the store manager has exerted sufficient operational and marketing efforts.

Specifically, we find an optimal compensation plan for the multitasking store manager, under the monotone likelihood ratio property (MLRP) that is commonly assumed in the principal-agent theory literature (see Laffont and Martimort 2009, pp. 164–165), consists of a base salary and a bonus paid to the store manager when either (i) inventory does not clear and the sales quantity exceeds an inventory-dependent threshold or (ii) inventory clears and the realized inventory level exceeds a threshold. Jointly, the two cases characterize the region of sales and inventory outcomes where the store manager receives a bonus. We show that, in the case where there are two effort levels in both operations and marketing, the bonus region exhibits a "mast and sail" structure, as illustrated in Figure 1. The "mast" corresponds to scenarios where inventory clears. The "sail" corresponds to scenarios where inventory does not clear.
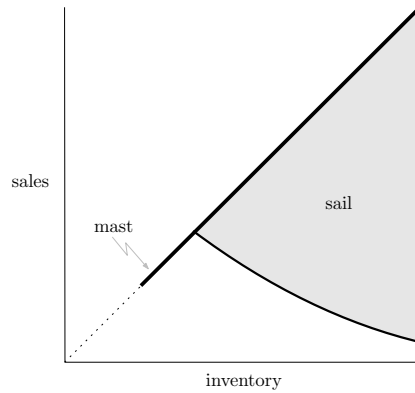

Figure 1   **The "mast-and-sail" bonus region of the optimal compensation plan under the MLRP assumption.**

Interestingly, the "mast and sail" structure of the bonus region gives rise to the nonmonotonicity of the optimal compensation plan. Given the same sales outcome, scenarios exist in which the store manager receives the bonus at some inventory level, but no longer so at a higher inventory level. In other words, the store manager may lose his bonus for achieving *better* inventory performance. To gain intuition about this nonmonotonicity result, note that when inventory is cleared, the realized demand is unobservable and capped by the inventory level. Thus, the firm's observed sales quantity is a *lower bound* of the realized demand. Given the same sales quantity, as inventory increases, the firm no longer clears the inventory. In this case, the observed sales quantity is *equal to* (as opposed to a lower bound of) the realized demand. An increased realized inventory may be informative of the store manager *not* exerting high marketing effort. This argument explains why the store manager

may seem to be penalized for a "better" inventory outcome. This phenomemon is a consequence of demand censoring.

These subtleties with nonmonotonicity underscore the inherent complexity of retail operations management and draw into stark relief the challenges that arise due to censored demand. Other difficulties for inventory management caused by demand censoring are well discussed in the literature (see, e.g., Feiler et al. 2012; Besbes and Muharremoglu 2013; Conlon and Mortimer 2013; Rudi and Drake 2014). This investigation provides yet another layer to this discussion, now in terms of designing incentives for store managers. Demand censoring distorts the informativeness of outcomes that can lead to nonmonotone compensation plans, adding further impetus for the design of accounting and physical tracking systems to better track demand in real time. Conlon and Mortimer (2013), for example, weigh the benefit of a wireless technology that provides detailed inventory information to help reduce the occurrence of demand censoring.

Additionally, under the MLRP assumption, the inventory-dependent sales threshold that warrants a bonus is decreasing in the inventory level.[3] Accordingly, sales and inventory outcomes act as substitutes in the optimal compensation plan once an inventory threshold has been met (precisely at the inventory level where the mast meets the sail in Figure 1). This result, seemingly inconsistent with the "weaker link" property of how demand and inventory drive sales, goes to the asymmetry of the operational and marketing outcomes signals—a low inventory not only reveals poor operational effort, but also reduces the principal's information by crippling the observability of true demand.

Under a slightly stronger version of the MLRP satisfied by many commonly considered distributions, likelihood ratios are *strictly* monotone. In this more stringent regime, we can additionally show that some intuitive compensation plans cannot be optimal. Indeed, one idea for a compensation plan with multiple signals is to give a bonus if each signal meets some minimum threshold. For example, a store manager receives a bonus if sales meet a quota and the available inventory level stays above some bar. We call such contracts "corner" compensation plans, because the two thresholds form a corner in the outcome space. The logic of corner compensation plans finds its trace in practice (e.g., Krishnan and Fisher 2005; DeHoratius and Raman 2007). Moreover, these plans are in line with known results in the single-tasking contract theory literature (e.g., Oyer 2000) where quota-bonus compensation plans are standard. However, we show that corner compensation plans are *not* optimal under the strict MLRP assumption.

---

[3] Throughout the paper, we use decreasing to mean "weakly" decreasing, in the sense that a function that is weakly decreasing may be constant on a subset of its domain. We say *strictly* decreasing when the function is decreasing but not constant. We use the same convention for the terminology increasing and strictly increasing.

Furthermore, through extensive numerical experiments, we exhibit typical cases in which corner compensation plans perform quite poorly in comparison to the optimal mast-and-sail contract and, in the worst case, performs arbitrarily bad. To develop contracts convenient to implement in practice, we approximate the optimal mast-and-sail contract by looking for simple contracts resembling it graphically. We find a type of monotone contract without a "mast" can perform well, but only if the "sail" properly reflects how inventory and sales act as complements or substitutes. By contrast, simple contracts with both a "mast" and a simple-to-compute linearized "sail" consistently achieve near-optimal performance.

The rest of the paper proceeds as follows. Section 2 reviews the literature. Section 3 formally states the multitasking store manager problem. Section 4 introduces a bang-bang approach for a broad class of incentive-design problems. Using this approach, Section 5 solves for the optimal compensation plan for the multitasking store manager and characterizes its structural properties. Section 6 derives managerial insights from the structure. Section 7 benchmarks the mast-and-sail contract versus simpler contracts analytically and numerically. Section 8 extends our model by relaxing some of the assumptions of the base model, including endogenizing starting inventory and the maximum level of store manager compensation. Section 9 concludes. All proofs are in the appendix. Finally, we provide further robustness checks and technical details in an online appendix.

2. Literature

Our paper contributes to three streams of literature: incentive design for store managers, salesforce compensation, and moral hazard in multitasking environments.

The retail operations literature has empirically documented the importance of incentive design. Most relevant to our paper is the work by DeHoratius and Raman (2007), who consider a store manager a *multitasking agent* who functions as both an inventory-shrinkage controller and a salesperson. DeHoratius and Raman (2007) substantiate the view that the store manager makes her effort decision across both job functions in response to incentives. In a detailed account of the day-to-day operations of a retail convenience chain, Krishnan and Fisher (2005) provide a process view of the range of a retail manager's responsibilities and detail the impact of incentive design on operational and marketing efforts, counting spoilage and shrinkage control as crucial areas over which store managers have substantial control. To the best of our knowledge, our paper is the first *analytical* treatment of optimal incentive design for a multitasking store manager. Our findings shed light on the nature of the relationship between marketing and operations, an issue that has inspired a voluminous literature (e.g., Shapiro 1977; Crittenden 1992; Ho and Tang 2004; Jerath et al. 2007).

Salesforce compensation has been studied extensively in the economics, marketing, and operations management literature (including, e.g., Basu et al. 1985; Lal and Srinivasan 1993; Raju and

Srinivasan 1996; Oyer 2000; Misra et al. 2004; Herweg et al. 2010; Jain 2012; Chen et al. 2018; Long and Nasiry 2018). The compensation literature has focused on two types of compensation plans—linear commission contracts and quota-bonus contracts. Both settings are well studied and have been applied in numerous contexts. However, the optimality of linear commission contracts has various caveats—its primary justification relies on a specific setting involving a normally distributed outcome and constant absolute risk-aversion (CARA) agent utility. By contrast, results on the optimality of quota-bonus contracts have followed from less restrictive conditions, namely, risk neutrality, limited liability, and general outcome distributions.[4] Drawing from the latter tradition, we establish the optimality of contracts in a multidimensional setting that are analogous to quota-bonus contracts in a single-dimensional setting.

Almost all of the salesforce compensation literature assumes unlimited inventory to meet demand generated by the salesperson. A recent stream of literature (Chu and Lai 2013; Dai and Jerath 2013, 2016) relaxes this assumption by incorporating demand censoring due to limited inventory. We study the compensation of a store manager who undertakes operational effort to increase the realized available inventory level, in addition to marketing effort to influence demand. By incorporating multitasking, our model differs significantly from formulations used in the salesforce compensation literature. As a result, our optimal compensation plan exhibits a structure that does not generalize the well-studied quota-bonus compensation plan (i.e., the salesperson receives a bonus for meeting a sales quota) in the single-tasking setting. Indeed, we show in Section 7 that several "intuitive" generalizations, including corner compensation plans (i.e., the store manager receives a bonus for meeting both sales and inventory thresholds), are not optimal and can perform poorly relative to the optimal compensation plan.

Our paper also relates to the accounting and economics literature (e.g., Holmstrom and Milgrom 1991; Feltham and Xie 1994; Dewatripont et al. 1999; Datar et al. 2001) on multitasking. In their seminal paper, Holmstrom and Milgrom (1991) model a multitasking agent whose job consists of multiple, concurrent activities, which jointly produce a multidimensional output signal. They focus on a linear compensation scheme and show that varying the weights of the compensation plan elicits changes in the agent's effort allocation. Our paper represents a departure from the multitasking literature in several meaningful ways. First, we consider a multidimensional output signal that affects the principal's utility in a *nonlinear* fashion (i.e., the sales quantity is equal to the minimum of demand and inventory), whereas, in the literature, the principal cares about an aggregated performance that is a linear combination of such dimensions. Second, our paper derives the *optimal* compensation plan, whereas most of the multitasking literature (following Holmstrom

---

[4] Limited liability captures an agent's aversion to downside risk and can be viewed as a type of risk aversion.

and Milgrom (1991)) assumes a linear compensation scheme without accounting for the form of the optimal compensation plan. Third, the observability of our multidimensional output signal is *imperfect* and microfounded through demand censoring. This observability issue has rich managerial implications. By comparison, the multitasking literature typically assumes perfect observability on all dimensions of the output signal.

In addition to contributing to the above three streams of literature, methodologically, our paper advances the moral hazard principal-agent theory by using a "bang-bang" approach to solve risk-neutral, limited-liability moral hazard problems with finitely many actions. Although "bang-bang" optimal control is a classical tool in economics, marketing, and operations (see, e.g., Sethi and Thompson 2000), to our knowledge, the application of this type of logic in the moral hazard literature is limited. We model the risk-neutral setting where the moral hazard problem is linear, and thus our characterizations are based on *extremal* solutions with a bang-bang structure. We explore this approach generally (i.e., beyond the store-manager context) to provide a methodological understanding of the approach, which we believe may be of separate interest for applied contract theory researchers. Our work applies result from a particularly cogent presentation of optimization in L_∞ spaces in Barvinok (2002, Sections III.5 and IV.12). This general setup treats linear optimal control as a special case.

Lastly, the "weaker link" property of the operational and marketing efforts plays a significant role in our analysis. The operations management and accounting literature (e.g., Chao et al. 2009; Baiman et al. 2010; Krishnan and Winter 2012, Section 8.1; Nikoofal and Gümüş 2018) has studied several similar settings where the outcome of a product is determined by the weakest of its several components; in particular, the familiar newsvendor model naturally gives rise to the "weaker link" property and can impact marketing decisions in meaningful ways (see, e.g., Jerath et al. 2017).

3.   Model

Consider a multitasking store manager (the agent) hired by a firm (the principal) to make operational effort e_o and marketing effort e_m. We assume that e_o and e_m take on at most finitely many values. In fact, our main structural results are for the case in which e_o and e_m may each take one of two different values ("high" and "low"). The reasons for these restrictions will become apparent in later sections. The principal cannot directly observe the effort choices of the store manager. Operational effort concerns increasing available inventory, and marketing effort concerns increasing sales. We assume both the realized inventory and demand levels are uncertain and that the store manager's efforts cannot be directly inferred by observing inventory and sales realizations.

Let us be more precise about the mechanics of operational effort and realized inventory. The firm supplies the store manager with an initial inventory level of $\bar{I}$. The realized available inventory level

$I \leq \bar{I}$ is all that is available to meet demand. The difference, $\bar{I} - I$, is not available to meet demand due to a variety of factors, as discussed in the introduction. Operational effort stochastically affects these factors to improve realized inventory. In our main analysis, we assume the initial inventory level $\bar{I}$ is exogenously given and focus our attention on the underlying incentive issues for effectively handling a given stock of inventory. Later, in Section 8.2, we endogenize $\bar{I}$ as a decision of the principal. We never consider $\bar{I}$ to be a choice of the store manager. In particular, our notion of operational effort *does not* entail forecasting or planning efforts for choosing a better starting initial inventory level $\bar{I}$.

We denote the cumulative distribution function of realized available inventory I by $F(i|e_o)$ and its probability density function by $f(i|e_o)$, where $i \in [0, \bar{I}]$. Similarly, we denote demand by Q, its cumulative distribution function by $G(q|e_m)$, and its probability density function by $g(q|e_m)$ for $q \in [0, \bar{Q}]$, where $\bar{Q}$ is an upper bound on demand.[5] These assumptions on the distribution of realized inventory and demand imply operational effort does not affect sales and marketing effort does not impact the realization of available inventory. Accordingly, for every effort level, the random variables I and Q are independent.

Regarding the density functions f and g, we assume they are both continuous functions of their first argument. Moreover, we assume the output distributions $f(I|e_o)$ and $g(Q|e_m)$ satisfy the monotone likelihood ratio property (MLRP); that is,

$$\frac{f(i|e_o)}{f(i|\hat{e}_o)} \text{ decreasing in } i \text{ for } e_o < \hat{e}_o \text{ and } \frac{g(s|e_m)}{g(s|\hat{e}_m)} \text{ decreasing in } s \text{ for } e_m < \hat{e}_m. \tag{1}$$

As a standard assumption in the literature (see, e.g., Grossman and Hart 1983; Rogerson 1985), the MLRP implies that a better inventory (demand) outcome is more informative of the fact that the store manager has exerted operational (marketing) effort. The MLRP is satisfied by most of the commonly-used families of distributions.

The (random) sales outcome is denoted $S \triangleq \min\{I, Q\}$. To reflect the phenomenon of demand censoring, we assume both the firm and store manager can observe the realized inventory level and the sales outcome, but neither can observe the realized demand in excess of the realized inventory level. We assume $\bar{Q} \geq \bar{I}$ to allow for the possibility that demand is censored at its highest level.

The store manager is effort averse. Her disutility from exerting efforts (e_o, e_m) is given by $c(e_o, e_m)$. We assume $c(e_o, e_m)$ is increasing in both dimensions of effort.

The firm's problem is to design a compensation plan $w(I, S)$ to maximize its total expected revenue less the total expected compensation to the store manager. We assume both firm and store

---

[5] Following the vast majority of the moral hazard literature, we assume a constant support for both demand and inventory outcomes. If either outcome has a support that moves with effort, the well-known "Mirrlees argument" applies: The firm can detect, with a positive likelihood, that the store manager has deviated from the desired action (Mirrlees 1999).

manager are risk neutral but with limited liability, bounding w below by $\underline{w}$ and above by $\bar{w}$. The lower bound on compensation ($\underline{w}$) is normalized to zero without loss. The latter ($\bar{w}$) implies the firm is budget constrained and cannot compensate beyond $\bar{w}$. This budget $\bar{w}$ is known to both the firm and the store manager. Assuming an upper bound for the compensation level is fairly common in the contract theory literature (e.g., Holmstrom 1979; Page 1987; Innes 1990; Arya et al. 2007; Jewitt 2008; Bond and Gomes 2009). In particular, Bond and Gomes (2009, p. 177) provide a variety of motivations for it, such as "a desire to limit the pay of an employee to less than his/her supervisor." In our main analysis, we continue to assume $\bar{w}$ is given, but later, in Section 8.1, we generalize the upper bound on $w(i, s)$ to be a more general resource constraint that is an integrable function of i and s.

The sequence of events is as follows. First, the firm offers a compensation plan $w(i, s)$, as a function of all possible inventory and sales levels, to the store manager who either takes it or leaves it. Second, if the store manager accepts the compensation plan, the store manager chooses an operational effort e_o and a marketing effort e_m; both efforts are exerted simultaneously. Third, both inventory I and demand Q outcomes are realized and inventory and sales $S = \min\{Q, I\}$ are observed. Each unit of met demand garners the principal a margin of r, unmet demand is lost and unobserved, and unused inventory is salvaged at a return normalized to zero. Fourth, the firm compensates the store manager according to $w(\cdot, \cdot)$. Because initial inventory $\bar{I}$ is given, the cost of procuring inventory is sunk. Accordingly, we may formulate the firm's problem as

$$\max_{w, e_o^*, e_m^*} \quad r\mathbb{E}[S|e_o^*, e_m^*] - \mathbb{E}[w(I, S)|e_o^*, e_m^*] \tag{2a}$$

$$\text{subject to} \quad S = \min\{Q, I\} \tag{2b}$$

$$\mathbb{E}[w(I, S)|e_o^*, e_m^*] - c(e_o^*, e_m^*) \geq \underline{U} \tag{2c}$$

$$\mathbb{E}[w(I, S)|e_o^*, e_m^*] - \mathbb{E}[w(I, S)|e_o, e_m] \geq c(e_o^*, e_m^*) - c(e_o, e_m) \text{ for all } (e_o, e_m) \tag{2d}$$

$$0 \leq w(i, s) \leq \bar{w} \text{ for all } (i, s), \tag{2e}$$

where the expectation $\mathbb{E}[\cdot|e_o, e_m]$ is taken over the joint distribution of I and S at the specified effort levels e_o and e_m. The participation constraint (2c) ensures the store manager's expected net payoff is no lower than a given reservation utility $\underline{U}$, and the store manager's incentive compatibility constraint (2d) ensures choosing (e_o^*, e_m^*) over all other effort levels is optimal for the store manager.

To reflect our primary motivating context, we refer to problem (2) as a *multitasking store manager* problem, although the same framework can be applied to a variety of situations involving multitasking between operational and marketing activities. This problem is conceptually challenging. Indeed, it is a bilevel optimization problem with an *infinite*-dimensional decision variable w. Deriving the form of an optimal compensation plan $w(\cdot, \cdot)$ requires a methodical exploration of

optimality conditions in this setting. Analyzing a general moral hazard problem is a Herculean task (if not intractable when few assumptions are made). Fortunately, the multitasking store manager problem has a compelling structure allowing for thorough analysis. We undertake this task in the next section, where we introduce a bang-bang approach to analyze a structured class of moral hazard problems that contains (2) as a special case.

4. A Bang-Bang Approach

In this section, we develop a method to study a class of risk-neutral moral hazard problems with finite agent action sets. We believe this method has general applicability beyond the multitasking store manager setting, and so we describe the approach in a neutral notation not overly specific to its use in this paper.

Consider a moral-hazard problem between one principal and one agent. The agent has a finite set of actions $\mathcal{A} = \{\vec{a}^1, \vec{a}^2, \ldots, \vec{a}^m\}$ from which to choose (we use the "arrow" notation $\vec{a}$ to denote a vector). In the multitasking setting, this assumption implies a finite number of operational effort levels e_o and a finite number of marketing effort levels e_m exist and each action $\vec{a} \in \mathcal{A}$ is a *pair* of efforts $\vec{a} = (e_o, e_m)$.

The agent incurs a cost $c(\vec{a})$ for taking action $\vec{a} \in A$, where we assume $c(\vec{a})$ is increasing in $\vec{a}$. The output is a vector $\vec{x} \in \mathcal{X}$, where $\mathcal{X}$ is a compact subset of $\mathbb{R}^n$, for some integer n.[6] That is, we consider the possibility of multiple signals. The random output X has density function $f(\vec{x}|\vec{a})$, where $f(\cdot|\vec{a})$ is in $L^1(\mathcal{X})$[7] for all $\vec{a} \in \mathcal{A}$ and $f(\vec{x}|\vec{a}) > 0$ for all $\vec{x} \in \mathcal{X}$ and $\vec{a} \in \mathcal{A}$. This general formulation allows for the possibility that the signals are correlated and depend on combinations of efforts.

The principal offers the agent the wage contract $w : \mathcal{X} \to \mathbb{R}$ that pays out according to the realized outcome. The principal values output $\vec{x} \in \mathcal{X}$ according to the valuation function $\pi : \mathcal{X} \to \mathbb{R}$. The agent has limited liability and must receive a minimum wage of $\underline{w}$ almost surely. We normalize $\underline{w}$ to zero. Moreover, the principal has a constraint that tops compensation out at $\bar{w}$; that is, $w(x) \leq \bar{w}$ for almost all $x \in \mathcal{X}$. Finally, the agent has a reservation utility $\underline{U}$ for her next-best alternative.

Both principal and agent are assumed to be risk neutral. The expected utility of the principal is denoted $V(w, \vec{a}) \triangleq \int_{x \in \mathcal{X}} (\pi(\vec{x}) - w(\vec{x})) f(\vec{x}|\vec{a}) dx$ and the expected utility of the agent is $U(w, \vec{a}) \triangleq \int_{\vec{x} \in \mathcal{X}} w(\vec{x}) f(\vec{x}|\vec{a}) d\vec{x} - c(\vec{a})$. The moral-hazard problem is

$$\max_{w, \vec{a}} \quad V(w, \vec{a}) \tag{3a}$$

---

[6] The compactness condition of $\mathcal{X}$ is not overly restrictive. If the original space of signals is unbounded, for instance, a transformation of the signal could make the signal space compact. For instance, tasking the transformation $\frac{e^x}{1+e^x}$ of the original signal x, in each dimension, can achieve the desired goal.

[7] The notation $L^1(\mathcal{X})$ denotes the space of all absolutely integrable functions on $\mathcal{X}$ with respect to Lebesgue measure on $\mathbb{R}^n$.

$$\text{subject to} \quad U(w,\vec{a}) \geq \underline{U} \tag{3b}$$

$$U(w,\vec{a}) - U(w,\vec{a}^i) \geq 0 \text{ for } i = 1,2,\dots,m \tag{3c}$$

$$0 \leq w \leq \overline{w}. \tag{3d}$$

The single constraint $U(w,\vec{a}) - U(w,\vec{a}^i) \geq 0$ for some $i \in \{1,2,\dots,m\}$ is called a *no-jump (NJ)* constraint. The incentive-compatibility constraint (3c) is the intersection of all no-jump constraints.

Following the two-step solution approach developed by Grossman and Hart (1983), we suppose an implementable target action $\vec{a}^*$ has been identified. This approach reduces the problem to

$$\min_{w} \quad \int_{\vec{x} \in \mathcal{X}} w(\vec{x}) f(\vec{x}|\vec{a}^*) d\vec{x} \tag{4a}$$

$$\text{subject to} \quad \int_{\vec{x} \in \mathcal{X}} w(\vec{x}) f(\vec{x}|\vec{a}^*) d\vec{x} \geq \underline{U} \tag{4b}$$

$$\int_{\vec{x} \in \mathcal{X}} R_i(\vec{x}) w(\vec{x}) f(\vec{x}|\vec{a}^*) d\vec{x} \geq c(\vec{a}^*) - c(\vec{a}^i) \text{ for } i \in \{1,2,\dots,m\} \text{ such that } \vec{a}^i \neq \vec{a}^* \tag{4c}$$

$$0 \leq w \leq \overline{w}, \tag{4d}$$

where we use the fact that $V(w,\vec{a}) = \mathbb{E}[\pi(\vec{x})|\vec{a}^*] - \int_{\vec{x} \in \mathcal{X}} w(\vec{x}) f(\vec{x}|\vec{a}^*) d\vec{x}$, drop the constant $\mathbb{E}[\pi(\vec{x})|\vec{a}^*]$ from the objective, convert to a minimization problem, and simplify the constraint (3c) using the definition

$$R_i(\vec{x}) \triangleq 1 - \frac{f(\vec{x}|\vec{a}^i)}{f(\vec{x}|\vec{a}^*)} \tag{5}$$

for $i = 1,2,\dots,m$. Finally, we drop the no-jump constraint for $\vec{a}^i = \vec{a}^*$, because this constraint is always satisfied with equality.

A *bang-bang contract* is a feasible solution w to (4), where $w(\vec{x}) \in \{0,\overline{w}\}$ for almost all $\vec{x} \in \mathcal{X}$. The goal of the next few results is to establish that an optimal bang-bang contract exists. First, we show an optimal contract exists:

LEMMA 1. *An optimal contract for* (4) *exists.*

Let W denote the set of feasible contracts to (4). An *extremal contract* of W is a contract that cannot be written as the convex combination of two other feasible contracts. That is, $w \in W$ is an extremal contract if there does not exist $w^1, w^2 \in W$ and $\lambda_1, \lambda_2 \in (0,1)$ with $\lambda_1 + \lambda_2 = 1$ such that $w = \lambda_1 w^1 + \lambda_2 w^2$. The next result is a consequence of Barvinok (2002, Proposition III.5.3).

LEMMA 2. *Every extremal feasible contract to* (4) *is a bang-bang contract.*

The previous two lemmas, along with Bauer's Maximum Principle (Aliprantis and Border 2006, Theorem 7.69), allow us to show that an optimal bang-bang contract exists.

THEOREM 1. *An optimal bang-bang contract for* (4) *exists.*

Our next goal is to characterize when an optimal bang-bang contract takes the value of 0 and when it takes the value $\bar{w}$. We show this characterization is associated with a *trigger* value of a weighted sum of appropriately defined covariances of the contract with the likelihoods of outcomes under different actions. The analysis here builds on Barvinok (2002, Section IV.12). The results of that section are established for linear programs over $L^\infty[0,1]$, where feasible solutions are bounded to be between 0 and 1. It is straightforward to extend these results to problems over compact set $\mathcal{X}$, where contracts are bounded between 0 and $\bar{w}$. Details of this adaptation are omitted. The result of this analysis yields the following result.

THEOREM 2. *There exist nonnegative multipliers ω_i and a "target" t such that there is an optimal solution to* (4) *of the form:*[8]

$$w^*(\vec{x}) = \begin{cases} \overline{w} & if \ \sum_{i=1}^m \omega_i R_i(\vec{x}) \geq t. \\ 0 & otherwise. \end{cases} \tag{6}$$

where $\sum_{i=1}^m \omega_i R_i(\vec{x}) = 1$ holds.

Let $B \triangleq \{\vec{x} \in \mathcal{X} : \sum_{i=1}^m \omega_i R_i(\vec{x}) \geq t\}$ denote the *bonus region* of the compensation plan w^*. In other words, $w^*(\vec{x})$ evaluates to $\bar{w}$ inside B and zero outside B.

The contract in (6) has a compelling economic interpretation. Consider the condition

$$\sum_{i=1}^m \omega_i R_i(\vec{x}) \geq t \tag{7}$$

that defines the bonus region B. Because the ω_i are nonnegative and sum to one, the left-hand side is a weighted sum of likelihood ratios that can be viewed as a measure of the information value (or informativeness) of outcome $\vec{x}$ for determining if the agent took the target action $\vec{a}^*$. For the given outcome $\vec{x}$, larger values of $R_i(\vec{x})$ are associated with actions $\vec{a}^i$ where the outcome $\vec{x}$ is less likely under action $\vec{a}^i$ than action $\vec{a}^*$. Thus, the larger $\sum_{i=1}^m \omega_i R_i(\vec{x})$ is, the less likely the agent is to have deviated from $\vec{a}^*$. The trigger condition (7) rewards outcomes whose informativeness exceeds the given threshold t. The weights ω_i fine tune how we measure this informativeness and are determined through solving a dual problem that "prices" the significance of deviations to different actions.

In light of this logic, we refer to contracts of the form (6) as *information-trigger* contracts (or simply *trigger* contracts). If the information value (as measured by $\sum_{i=1}^m \omega_i R_i(\vec{x})$) of a given outcome exceeds some trigger value, the agent is rewarded for that outcome.

---

[8] This assumes the function $\sum_{i=1}^m \omega_i R_i(\vec{x})$ has zero mass at the cutoff t. If there is positive mass at the cutoff, a lottery with payouts on 0 and $\bar{w}$ can characterize an optimal contract. We assume zero mass at the cutoff to avoid this additional complication.

The proof of Theorem 2 derives ω_i and t from solving a dual optimization problem. However, another approach is to solve a restricted class of the primal moral hazard problem (4), where contracts are information-trigger contracts of the form (6). If w is an information-trigger contract,

$$V(w, \vec{a}^*) = \bar{w} \int_{\substack{\vec{x} \in \mathcal{X} \text{ subject to} \\ \sum_{i=1}^{m} \omega_i R_i(\vec{x}) \geq t}} f(\vec{x}|\vec{a}^*) d\vec{x} = \bar{w} \mathbb{P}\left[\sum_{i=1}^{m} \omega_i R_i(\vec{X}) \geq t\right]$$

and

$$\int_{\substack{\vec{x} \in \mathcal{X} \text{ subject to} \\ \sum_{i=1}^{m} \omega_i R_i(\vec{x}) \geq t}} R_i(\vec{x}) w(\vec{x}) f(\vec{x}|\vec{a}) d\vec{x} = \bar{w} \mathbb{E}\left[R_i(\vec{X}) \middle| \sum_{i=1}^{m} \omega_i R_i(\vec{X}) \geq t\right],$$

where $\mathbb{P}[\cdot]$ is the probability measure and $\mathbb{E}[\cdot]$ is the expectation operator associated with $f(\cdot|\vec{a}^*)$. Using this notation, the restriction of (4) over trigger contracts of the form (6) is

$$\min_{\omega, t} \quad \bar{w}\mathbb{P}\left[\sum_{i=1}^{m} \omega_i R_i(\vec{X}) \geq t\right] \tag{8a}$$

$$\text{subject to} \quad \bar{w}\mathbb{P}\left[\sum_{i=1}^{m} \omega_i R_i(\vec{X}) \geq t\right] \geq \underline{U} \tag{8b}$$

$$\bar{w}\mathbb{E}\left[R_i(\vec{X}) \middle| \sum_{i=1}^{m} \omega_i R_i(\vec{X}) \geq t\right] \geq c(\vec{a}^*) - c(\vec{a}^i) \text{ for } i \in \{1, 2, \ldots, m\} \tag{8c}$$

$$\sum_{i=1}^{m} \omega_i = 1 \tag{8d}$$

$$\omega_i \geq 0 \text{ for all } i \in \{1, 2, \ldots, m\}. \tag{8e}$$

The next result relates optimality in this problem to the original problem (4).

THEOREM 3. *Problem* (8) *has the same optimal value as* (4). *Moreover, an optimal solution to* (8) *corresponds to an optimal solution to* (4).

Theorem 3 says that it suffices to solve the finite-dimensional problem (8) to solve the original moral hazard problem.

5.   Analysis

We now return to the multitasking store manager problem. First, in Section 5.1, we apply bang-bang control approach (Section 4) to this setting. Next, in Section 5.2, we explore the implications of this approach and illustrate the optimality of a mast-and-sail structure for the bonus region (captured in Figure 5(a) and as introduced in Figure 1 of the introduction). Section OA.1 provides a concrete numerical example to illustrate our development in this section.
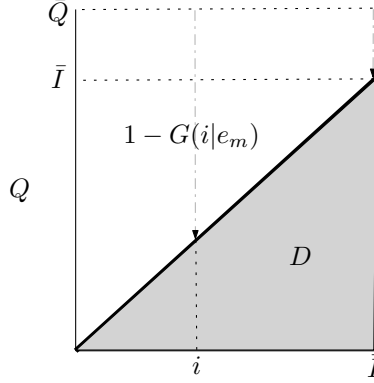
Figure 2   The domain D **(in gray) of the compensation plan** $w(I, S)$. **Note the thick line along the diagonal represents an accumulation of probability mass when demand exceeds inventory (the accumulation is represented by the downward-facing line from** $\bar{I}$ **to the** $45°$ **line in the figure).**

5.1. Optimal Compensation Plan

A critical object needed to define information-trigger compensation plans is the joint distribution of the outcome signals S and I, which is needed to define the analog to the ratio functions $R_i(x)$ that appear prominently in (6). Recall that demand Q and inventory I are assumed to be independent, and hence deriving their joint distribution is straightforward. A more difficult task is to derive the joint distribution of the sales S ($= \min\{Q, I\}$) and inventory I, which can be either independent of (when $Q < I$) or equal to each other (when $Q \geq I$). The following lemma provides the joint cumulative distribution function $\Pr(I \leq i, S \leq s | e_o, e_m)$ of I and S.

LEMMA 3. *The joint cumulative distribution function*

$$\Pr(I \leq i, S \leq s | e_o, e_m) = \begin{cases} F(s|e_o) + G(s|e_m)[F(i|e_o) - F(s|e_o)] & \text{if } s < i \\ F(i|e_o) & \text{if } s = i \end{cases}.$$

Lemma 3 shows that when $s = i$, the joint cumulative distribution function of S and I is the (marginal) distribution function of I. This is intuitive: we know $S \leq I$ and thus when $s = i$, the probability that S and I are both less than i is precisely the probability that I is less than i.

Before deriving the joint probability density function, we briefly discuss the domain of compensation plans in this setting. Note that

$$D \triangleq \big\{ (i, s) : 0 \leq s \leq i \text{ and } 0 \leq i \leq \bar{I} \big\}$$

is the domain of any feasible compensation plan because the firm cannot observe demand in excess of inventory. The domain D is shown in Figure 2. We also denote by

$$D^{\text{NSO}} \triangleq \{(i, s) \in D : s < i\} \text{ and } D^{\text{SO}} \triangleq \{(i, s) \in D : s = i\} \tag{9}$$

the regions of the domain where no stockout occurs and stockout occurs (respectively). This notation proves useful in succinctly describing our optimal compensation plans. For simplicity, we denote by $w(i)$ the compensation level when $s = i$; that is, we shorten $w(i, i)$ to $w(i)$.

The underlying measure of tuples (i, s) is absolutely continuous when $s < i$, whereas along the $45°$ line for each i, a point mass of weight $1 - G(i|e_m)$ at (i, i) is present. The joint probability density function of S and I is thus

$$h(i, s|e_o, e_m) = f(i|e_o)g(s|e_m) \tag{10}$$

for $s < i$, noting that $\frac{\partial^2}{\partial s \partial i}\big(F(s|e_o) + G(s|e_m)[F(i|e_o) - F(s|e_o)]\big) = f(i|e_o)g(s|e_m)$, and

$$h(i, i|e_o, e_m) = f(i|e_o)(1 - G(i|e_m)) \tag{11}$$

when $s = i$. We have from (9)–(11) that

$$\int_D h(i, s|e_o, e_m) \, di \, ds = \int_0^{\bar{Q}} \int_0^i g(s|e_m)f(i|e_o) \, ds \, di + \int_0^{\bar{I}}(1 - G(i|e_m)f(i|e_o) \, di$$

$$= \int_0^{\bar{I}} G(i|e_m)f(i|e_o) \, di + \int_0^{\bar{I}}(1 - G(i|e_m)f(i|e_o) \, di = 1,$$

and hence h is a legitimate probability density function.

Given this density function, using (5), we represent the ratio function $R_{e_o, e_m}(i, s)$ as

$$R_{e_o, e_m}(i, s) = 1 - \frac{\mathbb{I}[i > s]f(i|e_o)g(s|e_m) + \delta(i = s)f(i|e_o)(1 - G(i|e_m)}{\mathbb{I}[i > s]f(i|e_o^*)g(s|e_m^*) + \delta(i = s)f(i|e_o^*)(1 - G(i|e_m^*)},$$

where $\mathbb{I}[\cdot]$ is the indicator function and $\delta(i = s)$ is a Dirac function at i. We will describe an optimal information-trigger compensation plan in two different scenarios: (i) where $i > s$ (no stockout) and (ii) where $i = s$ (stockout), by defining appropriate ratio functions. In the no stockout (NSO) case,

$$R_{e_o, e_m}^{\text{NSO}}(i, s) = 1 - \frac{f(i|e_o)g(s|e_m)}{f(i|e_o^*)g(s|e_m^*)}, \tag{12}$$

and in the stockout (SO) case,

$$R_{e_o, e_m}^{\text{SO}}(i) = 1 - \frac{f(i|e_o)(1 - G(i|e_m)}{f(i|e_o^*)(1 - G(i|e_m^*)}. \tag{13}$$

Thus, by Theorem 2, an optimal compensation plan takes the following form:

$$w^*(i, s) = \begin{cases} w^{\text{NSO}}(i, s) & \text{if } (i, s) \in D^{\text{NSO}} \\ w^{\text{SO}}(i) & \text{if } (i, s) \in D^{\text{SO}} \end{cases}, \tag{14}$$

where

$$w^{\text{NSO}}(i, s) = \begin{cases} \overline{w} & \text{if } \sum_{e_o, e_m} \omega_{e_o, e_m} R_{e_o, e_m}^{\text{NSO}}(i, s) \geq t \\ 0 & \text{otherwise.} \end{cases}$$

and

$$w^{\mathrm{SO}}(i) = \begin{cases} \overline{w} & \text{if } \sum_{e_o,e_m} \omega_{e_o,e_m} R^{\mathrm{SO}}_{e_o,e_m}(i) \geq t \\ 0 & \text{otherwise,} \end{cases}$$

for some choice of t and nonnegative ω_{e_o,e_m} where $\sum_{e_o,e_m} \omega_{e_o,e_m} = 1$.

Recall that B denotes the bonus region of the information-trigger compensation plan w^* defined in (6). We adopt that notation here and refine it further by setting

$$B^{\mathrm{NSO}} \triangleq \left\{ (i,s) \in D^{\mathrm{NSO}} : \sum_{e_o,e_m} \omega_{e_o,e_m} R^{\mathrm{NSO}}_{e_o,e_m}(i,s) \geq t \right\} \tag{15}$$

and

$$B^{\mathrm{SO}} \triangleq \left\{ (i,s) \in D^{\mathrm{SO}} : \sum_{e_o,e_m} \omega_{e_o,e_m} R^{\mathrm{SO}}_{e_o,e_m}(i,s) \geq t \right\}. \tag{16}$$

A key observation here is that the bonus region B^{NSO} is possibly a full-dimensional subset of the nonstockout region of the domain D^{SO}, whereas the bonus region B^{SO} is a one-dimensional set along the 45° line D^{SO}. As we make concrete in the next sections, these two bonus regions need not "overlap," even when the output distributions have a great deal of regularity. This observation underscores the complexity of incentive design in this multitasking setting.

5.2. Structural Properties

In the rest of this section, we make more concrete the structure of the optimal compensation plan (14) in a special multitasking setting with two levels—"high" (H) and "low" (L)—for each of the operational and marketing efforts. In the notation of Section 4, $\mathcal{A} = \{(e_o^H, e_m^H), (e_o^H, e_m^L), (e_o^L, e_m^H), (e_o^L, e_m^L)\}$. We will also assume the target action is (e_o^H, e_m^H), that is, for the store manager to make her best effort in both operations and marketing. Recall from Section 3 that we assume both inventory and demand outcomes satisfy the MLRP. To explore the implications of the MLRP assumption, it is useful to represent $R^{\mathrm{NSO}}_{e_o,e_m}(i,s)$ and $R^{\mathrm{SO}}_{e_o,e_m}$ explicitly:

$$R^{\mathrm{NSO}}_{e_o^H,e_m^L}(i,s) = 1 - \frac{g(s|e_m^L)}{g(s|e_m^H)}, R^{\mathrm{NSO}}_{e_o^L,e_m^H}(i,s) = 1 - \frac{f(i|e_o^L)}{f(i|e_o^H)}, \text{ and } R^{\mathrm{NSO}}_{e_o^L,e_m^L}(i,s) = 1 - \frac{f(i|e_o^L)g(s|e_m^L)}{f(i|e_o^H)g(s|e_m^H)}; \tag{17}$$

$$R^{\mathrm{SO}}_{e_o^H,e_m^L}(i) = 1 - \frac{1-G(i|e_m^L)}{1-G(i|e_m^H)}, R^{\mathrm{SO}}_{e_o^L,e_m^H}(i) = 1 - \frac{f(i|e_o^L)}{f(i|e_o^H)}, \text{ and } R^{\mathrm{SO}}_{e_o^L,e_m^L}(i) = 1 - \frac{f(i|e_o^L)(1-G(i|e_m^L))}{f(i|e_o^H)(1-G(i|e_m^H))}. \tag{18}$$

Note, in particular, that $R^{\mathrm{NSO}}_{e_o^H,e_m^L}(i,s)$ is constant in i and $R^{\mathrm{NSO}}_{e_o^L,e_m^H}(i,s)$ is constant in s. This observation has important consequences for later development. The MLRP lends monotonicity to the ratios on the right-hand sides of (12) and (13) that allows us to get a clearer picture of the structure of the bonus regions B^{NSO} and B^{SO} defined in (15) and (16), respectively. In addition, the definition of the MLRP (1) has implications for the ratios.

LEMMA 4. *(i) Both $\frac{f(i|e_o^L)}{f(i|e_o^H)}$ and $\frac{1-G(i|e_m^L)}{1-G(i|e_m^L)}$ are decreasing and nonconstant in i;*
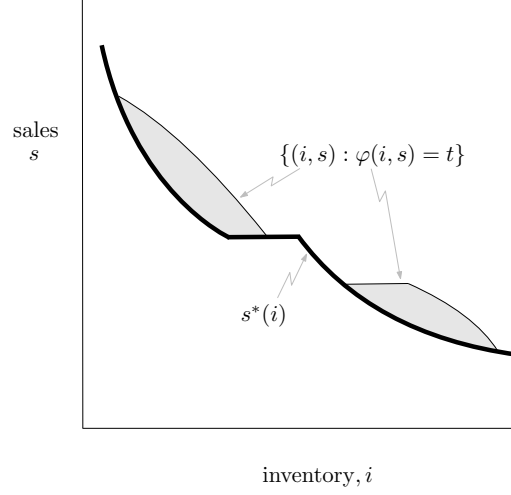
Figure 3   **Illustration of the level set** $\{(i,s): \varphi^{\mathsf{NSO}}(i,s)=t\}$ **of the function** φ^{NSO} **(defined in** (21)**) and the function**
$s^*(i)$ **(defined in** (20)**.**

(ii) $\frac{g(s|e_m^L)}{g(s|e_m^H)}$ *is decreasing and nonconstant in* s.

We first look into the structure of B^{NSO}. This is captured in the following result.

PROPOSITION 1. *A decreasing and continuous function* s^* *and inventory value* $i_s \in (0, \bar{I}]$ *exist such that*

$$B^{NSO} = \{(i,s): i \geq i_s \ and \ s^*(i) \leq s < i\}. \tag{19}$$

Specifically, the function $s^*(\cdot)$ is defined according to

$$s^*(i) \triangleq \min\{s : \varphi^{\mathrm{NSO}}(i,s) = t\}, \tag{20}$$

where

$$\varphi^{NSO}(i,s) \triangleq \sum_{e_o,e_m} \omega_{e_o,e_m} R^{\mathrm{NSO}}(i,s). \tag{21}$$

These objects are illustrated in Figure 3. The critical inventory value i_s is the unique fixed point of s^*. That is, it is defined by

$$s^*(i_s) = i_s. \tag{22}$$

Because $s^*(i)$ is a decreasing and continuous function of i on its domain, the bonus region resembles the one shown in Figure 4(a). We call the shape of this region a "sail" for obvious reasons; accordingly, the subscript "s" in the inventory threshold i_s stands for "sail." When B^{NSO} has zero measure, set $i_s = \bar{I}$—its maximum possible value—and the "sail" disappears.

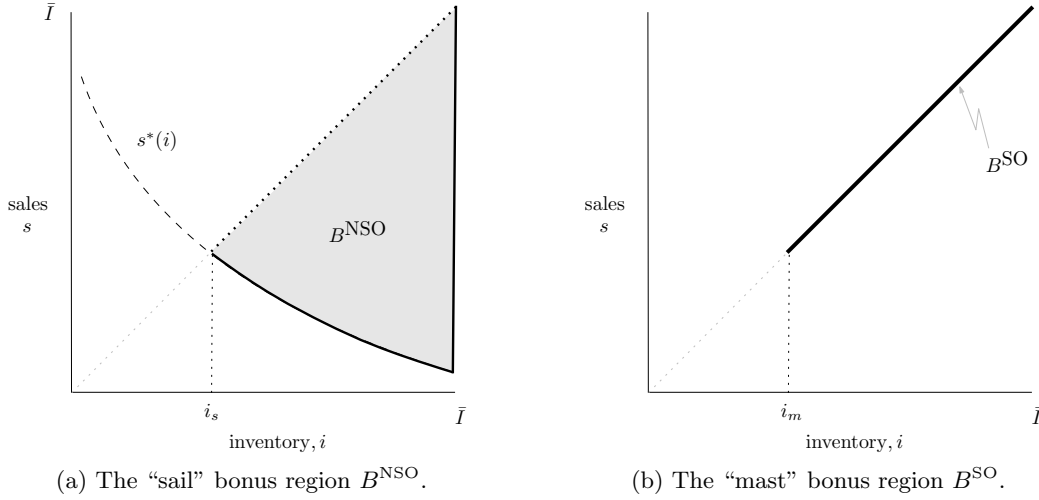We turn now to the stockout scenario and study the shape of B^{SO}.

(a) The "sail" bonus region B^{NSO}.

(b) The "mast" bonus region B^{SO}.

Figure 4   **Illustrations of the "sail" and "mast" regions.**

PROPOSITION 2. *An inventory value $i_m \in (0, \bar{I}]$ exists, such that*

$$B^{SO} = \{(i, s) : s = i \geq i_m\}. \tag{23}$$

When B^{SO} has a positive measure, we set

$$i_m = \min \left\{ i : \varphi^{\mathrm{SO}}(i) = t \right\},$$

where

$$\varphi^{\mathrm{SO}}(i) \triangleq \sum_{e_o, e_m} \omega_{e_o, e_m} R^{\mathrm{SO}}_{e_o, e_m}(i).$$

Again, the existence of such an i_m in this case follows from the monotonicity and continuity properties of the $R^{\mathrm{SO}}_{e_o, e_m}$ under the MLRP. If B^{SO} has zero measure, we simply set $i_m = \bar{I}$.

Figure 4(b) gives a visualization of the bonus region B^{SO}. We call this region a "mast" that is set up at a 45° angle. Accordingly, the subscript "m" of the inventory threshold i_m stands for "mast."

Taken together, the bonus region of the optimal compensation plan w^* defined in (14) is the union of the regions in (19) and (23). Figures 5(a) and 5(b) illustrate two of the possible structures of this union that result from the regions failing to "overlap" perfectly. When $i_s > i_m$, the bonus region has an inherent nonconvexity at (i_s, i_s), as illustrated in Figure 5(a). The shape in this figure makes clear our usage of the phrase "mast and sail" to describe the bonus region of an optimal compensation plan. When $i_s < i_m$, the "mast and sail" (with what looks like a mast that is too short for its sail) overlap to form a region that is not closed, as illustrated in Figure 5(b).[9]

---

[9] Indeed, when the MLRP fails and we only have first-order stochastic dominance in $g(\cdot | e_m)$, we demonstrate a numerical example whose bonus region is not closed (as in Figure 5(b)). This example is not included in the manuscript due to page limitations, and is available upon request.
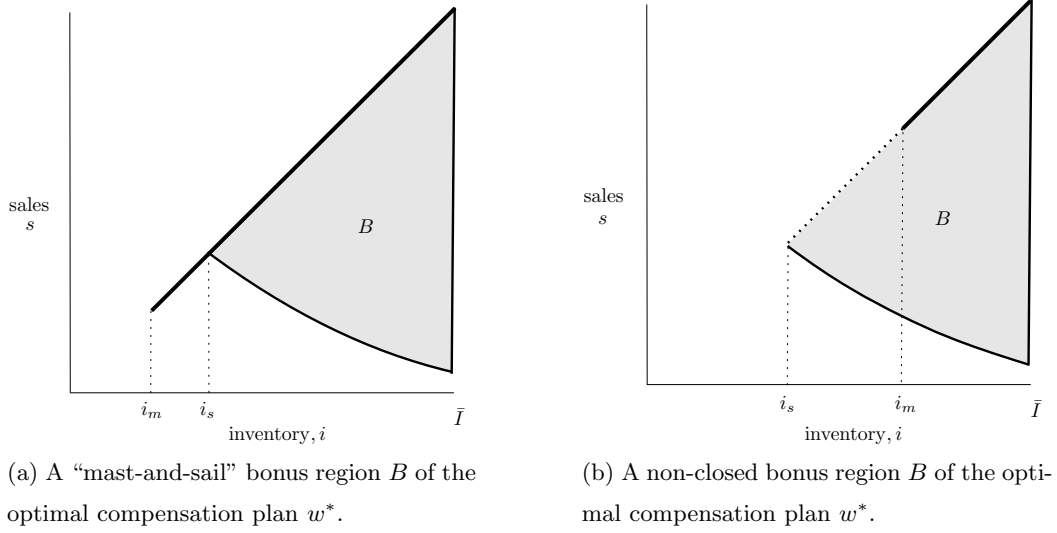
(a) A "mast-and-sail" bonus region B of the optimal compensation plan w^*.

(b) A non-closed bonus region B of the optimal compensation plan w^*.

Figure 5   **Two possible structures of the union of the "mast" and the "sail" bonus regions.**

PROPOSITION 3. *In every optimal compensation plan w^* of the form* (14), *we have $i_s^* \geq i_m^*$. Moreover, if $\min\{\omega_{e_o^H, e_m^L}, \omega_{e_o^H, e_m^L}\} > 0$, then $i_s^* > i_m^*$.*

The above proposition states that out of the two structures illustrated in Figure 5, only the "mast-and-sail" structure seen in Figure 5(a) is possible in the optimal compensation plan. This result leverages the fact that both $\varphi^{\mathrm{NSO}}(i_s, i_s)$ and $\varphi^{\mathrm{SO}}(i_m)$ are equal to t, and thus equal to each other. The equality $\varphi^{\mathrm{NSO}}(i_s, i_s) = \varphi^{\mathrm{SO}}(i_m)$ has implications for the relative sizes of i_s and i_m, in light of Lemma 4, that hold when both f and g satisfy the MLRP.

6. Implications for Incentive Design

Having established (in Proposition 3) that the optimal compensation plan has a "mast and sail" structure, we turn to reveal implications for incentive design that this structure reveals. We discuss the (non-)monotonicity of the compensation plan in Section 6.1 and the relationship between operations and marketing in the optimal compensation plan in Section 6.2.

6.1. (Non-)Monotonicity

We first explore the monotonicity of an optimal compensation plan whose bonus region has the mast-and-sail structure seen in Figure 5(a). Because a compensation plan has two arguments (i and s), more than one notion of monotonicity exists. First, we say $w(i, s)$ is *monotone in i* if $w(i', s) \leq w(i'', s)$ for every $(i', s), (i'', s) \in D$ and $i' \leq i''$. Similarly, we say $w(i, s)$ is *monotone in s* if $w(i, s') \leq w(i, s'')$ for every $(i, s'), (i, s'') \in D$ and $s' \leq s''$. Finally, we say $w(i, s)$ is (strictly) *jointly monotone* if $w(i', s') \leq w(i'', s'')$ for all $(i', s'), (i'', s'') \in D$ with $i' < i''$ and $s' < s''$.[10] Geometrically,

---

[10] We say "strictly" here because we require strict improvement in both the inventory and sales outcomes. Note that allowing $i' = i''$ or $s' = s''$ in the definition of joint monotonicity is a case that can be handled by one of the two
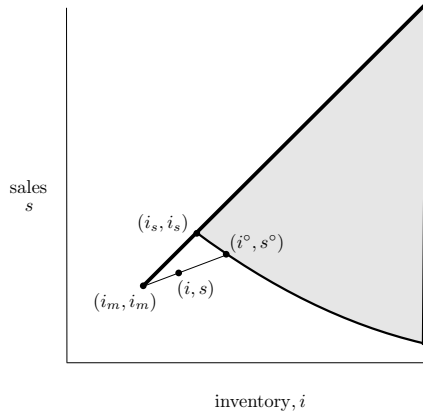
Figure 6   **An illustration of the lack of joint monotonicity of an optimal compensation plan** w^* **under the MLRP assumption.**

monotonicity in i calls for monotonicity along horizontal lines in Figure 5(a), monotonicity in s calls for monotonicity along vertical lines, and joint monotonicity calls for monotonicity when moving strictly up and to the right.

PROPOSITION 4. *For an optimal compensation plan* w^* *of the form* (14),

(i) w^* *is monotone in* s;

(ii) *if* $i_s = i_m$, *then* w^* *is also monotone in* i *and jointly monotone;*

(iii) *if* $i_s > i_m$, *then* w^* *is neither monotone in* i *nor jointly monotone.*

A careful investigation of Figure 6 provides the necessary intuition for understanding this result. Part (i) concerns monotonicity in the vertical direction, which clearly holds in the figure. Note that we never move beyond the 45° line in the vertical direction, and so clearly, once the store manager has a bonus for a given outcome (i, s), all higher sales at that inventory level also yield a bonus. Part (ii) concerns monotonicity in the horizontal direction. Moving a short distance horizontally from the bottom corner (i_m, i_m) of the mast drops the store manager's compensation from having the bonus to losing the bonus. Lastly, part (iii) concerns moving northeast in the graph. As shown in Figure 6, a move from the corner (i_m, i_m) of the mast to the point (i°, s°) where $i^\circ = i_s + \epsilon$ for some positive ϵ and $s^\circ = s^*(i^\circ)$ again drops the bonus for the store manager. This argument is carefully detailed in the proof of Proposition 4 in the appendix.

To say that an optimal compensation plan is not monotone in every sense is somewhat nonintuitive. Indeed, as seen in Figure 6, the store manager could be worse off for achieving strictly better inventory and sales outcomes. The source of nonmonotonicity comes from moving off the mast part of the bonus region and landing in the "gap" between the mast and sail. As briefly discussed in the

earlier definitions of monotonicity.

introduction, when stockout occurs along the mast part of the bonus region, the realized demand is censored by the inventory level. The firm's observed sales quantity is only a *lower bound* of the realized demand. The store manager might have made significant marketing effort that realized in a high demand level, which (possibly unluckily) available inventory was not able to meet. Given the same sales quantity, as inventory increases, the firm no longer experiences stockouts. In this case, the observed sales quantity is *equal to* (as opposed to a lower bound of) the realized demand. Thus, an increased realized inventory level may be informative of the fact that the store manager *has not* exerted high marketing effort. In other words, to encourage greater marketing effort, the firm is rewarding the possibility of a high demand realization when inventory stocks out. However, when the uncertainty surrounding realized demand (as opposed to sales) is removed, better performance is required to warrant a reward.

On a cautious note, nonmonotone contracts may give rise to perverse incentives if not handled properly. Suppose a store manager realizes inventory and sales (i', s') with $i' > i_m$ but does not receive a bonus. This scenario occurs, for instance, when $i' \in (i_m, i_s)$ and $i' < s' < s^*(i')$. If the store manager could hide some of the realized inventory (or claim it was "shrunk") to reveal an output of (i', i'), he would receive a bonus. This analysis, however, assumes the store manager can freely "shrink" inventory. In reality, such a practice is highly discouraged by the firm and efforts to monitor such behavior is discussed in DeHoratius and Raman (2007). When the chance that the store manager will be "caught" in the act of "shrinking" inventory, and penalized appropriately, is sufficiently high, nonmonotonicity of the compensation plan may persist in practice. We take up this issue in greater depth in Section 8.1.1, where we show how this type of thinking can translate into a natural upper bound on $\bar{w}$.

Note this nonmonotonicity may disappear if the lower part of the mast below the sail is removed from the bonus region (other ideas to recover monotonicity are explored in Section 7). However, no optimal compensation plan with this type of bonus region may exist. Such a change shifts effort away from marketing (which were previously rewarded when inventory is cleared) toward inventory. By only rewarding higher realized inventories, the store manager may need to sacrifice more in terms of marketing effort than is desirable for the firm. In other words, insisting on monotonicity of the optimal compensation plan may undervalue the store manager's marketing effortswhen stockouts occur.

Proposition 4 shows, at a high level, that demand censoring is tied to the issue of nonmonotonicity in store manager compensation. It represents another challenge associated with censoring in operations, which includes issues related to procurement, pricing, and forecasting (Perakis and Roels 2010; Huh et al. 2011; Besbes and Muharremoglu 2013; Jain et al. 2014). The "incentive challenge" associated with demand censoring, and its connection to nonmonotonicity in optimal

compensation, further underscores the operational benefits of being able to observe demand directly (and not just sales). In Section OA.4 of the online appendix, we look at the case in which demand is observable, and numerically evaluate the loss due to censoring in incentive design.

6.2. Operations and Marketing: Complements or Substitutes?

In the optimal compensation plan, should the operational and marketing performance metrics (i.e., inventory and demand) act as complements or substitutes? One might expect complementarity in the compensation plan because for the firm to benefit from an increase in inventory, the increase in inventory must be accompanied by an increase in demand. The next proposition, following from Proposition 1, suggests otherwise.

PROPOSITION 5. *As i increases,*

(i) *if $i_m \leq i < i_s$, the minimum sales quantity required for the store manager to qualify for the bonus strictly increases ("moving up the mast");*

(ii) *if $i \geq i_s$, the minimum sales quantity $s^*(i)$ required for the store manager to qualify for the bonus decreases ("slipping down the sail").*

Proposition 5(i) reveals that in the "mast" part of the bonus region (i.e., the region with $i < i_s$), echoing the "weaker link" property of our model setup, for the agent to receive the bonus, a high realized inventory level has to be accompanied by a high sales outcome. In other words, in the optimal compensation plan, operations and marketing act as complements when the observed inventory is not sufficiently high. The complementary in the compensation plan takes such an extreme form that the sales threshold is exactly equal to the inventory outcome. To understand why, note the "mast" part corresponds to stockout scenarios. Because the inventory is not sufficiently high, the firm expects the agent to generate a high-enough demand to clear all the inventory to demonstrate the agent has exerted a sufficient marketing effort.

In the "sail" part of the bonus region (i.e., the region with $i \geq i_s$), as the inventory level increases, the minimum sales quantity to receive the bonus *decreases*. One might expect a higher—not lower—sales threshold as the realized inventory level (i) increases. The intuition of this result is connected to the MLRP, which suggests the informative value of the observed signal (i, s) increases in both i and s. Note that the "sail" part corresponds to the scenarios without stockouts so that the true demand is equal to the observed sales quantity. Thus, the firm can infer the same likelihood that the store manager has exerted both operational and marketing efforts based on either (i) a low inventory level and a high demand outcome or (ii) a high inventory level and a lower demand outcome. For this reason, when the observed inventory outcome is sufficiently high, operations and marketing are substitutes in the optimal compensation plan.

Comparing Proposition 5(i) and (ii), we have the following implication for incentive design in a marketing-operations multitasking setting: when the observed inventory level is not sufficiently high, operations and marketing act as complements in the compensation plan, meaning the firm should expect a higher sales target when observing a higher inventory level. Once the observed inventory level exceeds a certain threshold, operations and marketing become substitutes in the compensation plan, meaning the firm should expect a lower sales target when observing a higher inventory level. The complementarity effect comes from demand censoring, whereas the substitute effect comes from the information trigger governing the optimal incentive design.

7.　Performance of Simple Compensation Plans

We have shown that an optimal mast-and-sail compensation plan exists under certain conditions. However, the optimal mast-and-sail structure remains somewhat nonintuitive and requires effort to compute (particularly in determining s^*). This motivates interest in exploring the performance of classes of "simple" compensation plans that have an intuitive structure and whose optimal parameters can be computed easily. The first class of simple compensation plans we explore includes bonus compensation plans whereby the agent receives a bonus if both a sales and inventory quota are met, and we term them "corner compensation plans" for reasons that we explore more carefully below. We also explore other simple compensation plans, inspired by the corner compensation plans and elements of the mast-and-sail compensation plan.

An interesting question is to assess the extent of optimality loss associated with simplicity, using the principal's profit under the optimal mast-and-sail compensation plan as the upper bound. We provide one analytical bound on the performance of corner contracts in a slightly restricted setting; namely, in Proposition 6 we show that corner compensation plans are never optimal in that setting.

Additional analytical bounds on the performance of simple compensation plans are hard to come by, in no small part because of the challenging nature of computing the precise structure of mast-and-sail compensation plans. The difficulty is that the weights ω_i and t must be computed to get a sense of the shape of the mast and sail. We believe that problem (8) and Theorem 3 provide our best hope for computing ω_i and t in general. However, (8) is a challenging optimization problem and, to our knowledge, does not readily admit analytical characterizations that can be used to provide bounds.

For the above reasons, numerical calculations are the primary standard we use to compare the various types of compensation plans.

7.1.　Corner Compensation Plans

A natural class of compensation plans to consider are those that offer a bonus if sales and inventory both meet some specified targets. These compensation plans build on the logic of the quota-bonus

contracts that are optimal in the risk-neutral setting in the salesforce compensation literature (Oyer 2000; Dai and Jerath 2013, 2016).

Two obvious generalizations of quota-bonus contracts arise in the multi-tasking setting. The first are compensation plans with some a and b where any outcomes (i,s) with $i \geq a$ *and* $s \geq b$ earn a bonus $0 \leq \beta \leq \bar{w}$. We call these plans *corner compensation plan* (a,b). See Figure 7(a) for an illustration. The second are compensation plans that allot bonuses when targets on the outcomes are rewarded individually, with what we call a *two-stage* bonus compensation plan. More precisely, any inventory level $i \geq a$ is rewarded β_i and any sales level $s \geq b$ is rewarded β_s. Any outcome (i,s) with $i \geq a$ *and* sales $s \geq b$ receives a combined bonus of $\beta_i + \beta_s$.

(a) The bonus region of a corner compensation plan.

(b) The bonus region(s) of a two-stage bonus compensation plan.
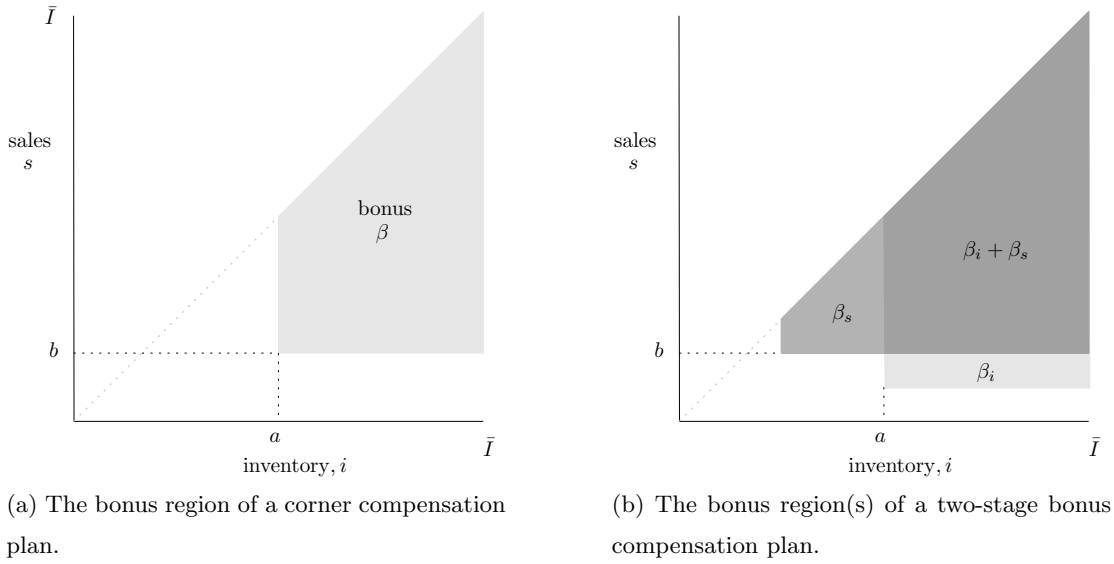
Figure 7   Illustrations of nonoptimal "intuitive" bonus regions.

The remaining goal in this subsection is to compare the performance of these compensation plans with the best mast-and-sail compensation plan. Because an optimal mast-and-sail compensation plan exists, it provides an upper bound on the performance of corner compensation plans. In fact, the next result shows that in particular cases, we know mast-and-sail compensation plans perform *strictly* better than the compensation plans featured in Figure 7.

We say f and g satisfy the *strict* MLRP; that is, (1) holds with weak inequalities replaced by strict inequalities. Many commonly studied families of distributions (e.g., binomial, exponential, log-normal, normal, and Poisson) satisfy the strict MLRP.

PROPOSITION 6. *Consider the multitasking store manager problem described in Section 5.2, with the further restriction that f and g satisfy the strict MLRP. An optimal compensation plan can be neither a two-stage bonus nor a corner compensation plan.*
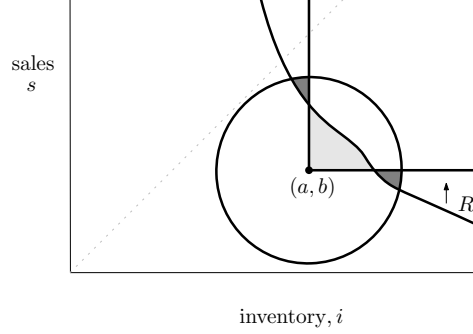
**Figure 8**   **Rounding the "corner" of a nonoptimal compensation plan.**

Assuming positive rents for the agent is common in the literature (e.g., Oyer 2000; Dai and Jerath 2013). The situation in which the agent earns no rents yields a first-best contract whereby the incentive issue does not have any "bite," and is thus less interesting as an incentive problem.

The proof of Proposition 6 is technical. Here, we provide its basic intuition. The essence of the proof is to show the bonus region of an optimal compensation plan cannot have a non-smooth "corner" like that seen at (a, b) in Figure 8. The proof shows such a bonus region can be "rounded" by removing a part of the bonus region (depicted in light gray) and adding new areas to the bonus region (depicted in dark gray). The claim is that this operation translates into a decrease in the expected payment to the store manager that nonetheless continues to satisfy incentive compatibility. Loosely speaking, the claim follows because, in the strict MLRP setting, the level sets of the weighted ratio function φ^{NSO} defined in (21) is a strictly decreasing function (whose graph is depicted as the curve R in Figure 8). Thus, an incentive-compatible compensation plan with a bonus region exhibiting a corner at (a, b) has a greater covariance with the weighted likelihood ratio in the dark-gray regions than the light-gray region. A compensation plan with the bonus region that swaps these regions must also satisfy incentive compatibility.

Using similar reasoning as the proof of this proposition, one can show the best corner compensation plan with bonus $\bar{w}$ outperforms every other corner compensation plan as well as the optimal two-stage bonus compensation plan. Accordingly, we focus our attention on optimal corner compensation plans with bonus $\bar{w}$. Moreover, observe that compensation plans rewarding sales *only* are achieved by setting $a = b$, and those rewarding inventory *only* are achieved by setting $b = 0$. Accordingly, single-tasking focused compensation plans are special cases of corner compensation plans and so are (weakly) dominated by the optimal corner compensation plan.[11]

We use numerical experiments to evaluate the performance of corner compensation plans and present two representative and *contrasting* scenarios in Figures 9(a) and 9(b), respectively. Figure 9(a) shows the performance of the corner compensation plan is nearly optimal when the

---

[11] See Section OA.2 for more discussion on sales-only contracts as a special case.
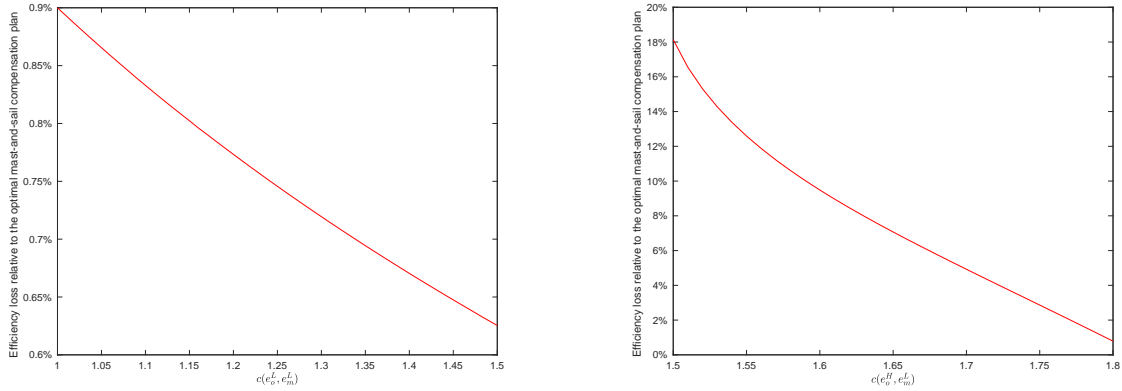
(a) Parameters: $\bar{w} = 10$, $c(e_o^H, e_m^L) = c(e_o^L, e_m^H) = 3.0$, and $c(e_o^H, e_m^H) = 3.5$. We vary $c(e_o^L, e_m^L)$ between 1 and 1.5.

(b) Parameters: $\bar{w} = 10$, $c(e_o^L, e_m^L) = 1$, $c(e_o^L, e_m^H) = 1.8$, and $c(e_o^H, e_m^H) = 3.5$. We vary $c(e_o^H, e_m^L)$ between 1.5 and 1.8.

Figure 9 **Performance of the optimal corner compensation plan (which is also the optimal quota-bonus compensation plan) relative to the optimal mast-and-sail compensation plan. We assume $F(i|e_o) = (H(i))^{e_o}$ and $G(s|e_m) = (L(s))^{e_m}$, where $H(i) = i$ and $L(s) = s$.**

marketing and operational activities are highly complementary in terms of the agent's cost structure (i.e., compared to exerting a high effort level in either operations or marketing, the marginal cost of exerting high effort levels in *both* marketing and operational activities is relatively low). By contrast, Figure 9(a) shows when the marketing and operational activities are not sufficiently complementary, the performance of the corner compensation plan is far from optimal.[12]

In certain cases, the corner compensation plan fails to induce the target action that is achievable under the optimal contract. In this case, although the corner compensation plan may lead to a lower expected compensation than the optimal compensation plan, the firm's expected sales quantity is also lower because of the store manager's lower effort than the desired one. Thus, under a sufficiently high unit revenue (so that the target action entails high effort in both operational and marketing activities), the firm's expected profit is higher under the mast-and-sail compensation plan. Indeed, for this type of scenario, we can show the efficiency loss under the corner compensation plan increases linearly in the unit revenue. We use an example to illustrate this result.

EXAMPLE 1. Consider the following instance where $e_o \in \{e_o^L, e_o^H\}$ and $e_m \in \{e_m^L, e_m^H\}$ where $e_o^L = e_m^L = 1$ and $e_o^H = e_m^H = 2$. The target action is $(e_o^H, e_m^H) = (2, 2)$. The cost function is $c(e_o^H, e_m^H) = 3.1$, $c(e_o^H, e_m^L) = 1$, $c(e_o^L, e_m^H) = 1.6$, and $c(e_o^L, e_m^L) = 0.1$. The resource constraint for the firm has $\bar{w} = 10$. For this instance, we can show the firm can use a mast-and-sail compensation plan with

---

[12] To validate the robustness of our numerical experiments, we have also conducted extensive numerical experiments using the power of other probability distributions for $H(i)$ and $L(s)$, including beta distribution, exponential distribution, gamma distribution, truncated normal distribution, normal distribution, and Weibull distribution; our findings are directionally the same.
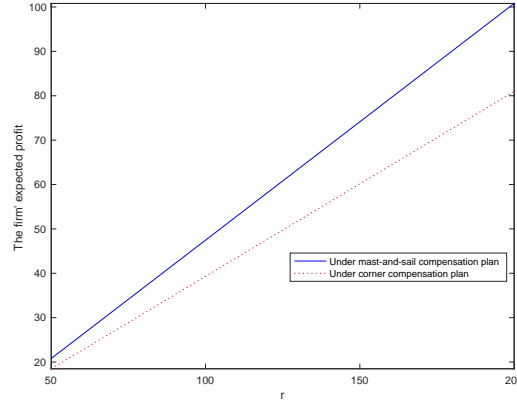
Figure 10   **The firm's expected profits under the mast-and-sail and corner compensation plans as functions of the per unit revenue rate r.**

$\omega^*_{e_o^L, e_m^L} = 0, \omega^*_{e_o^H, e_m^L} = 0.8602$, $\omega^*_{e_o^L, e_m^H} = 0.1398$, and $t^* = 0.1817$ to induce the target action, under which the store manager's probability of receiving the bonus is 58.70%. However, no corner contract exists that can induce the target action. Indeed, the best that the corner compensation plan can achieve is to induce (e_o^H, e_m^L) with parameters of $a^* = b^* = 0.6186$, under which the store manager's probability of receiving the bonus is 23.55%. We illustrate the firm's expected profits under both types of compensation plans in Figure 10, as a function of the per-unit revenue rate r.

7.2. Approximating Mast-and-Sail Contracts

So far, we have used a *single-tasking logic* to construct and evaluate corner compensation plans, with the sales-quota-bonus compensation plan being the simplest case, and found its performance depends on the store manager's cost structure and can be far from optimal. We now switch gears to using our mast-and-sail compensation plan as inspiration for designing simple contracts. Below, we examine two directions of approximating the mast-and-sail contract.

One direction is to connect the lower end of the mast and the lowest point of the sail. In this way, we create a type of contract with a "triangular sail," with two possible variants:

(a) *Weighted-sum threshold compensation plan*: This compensation plan determines whether to pay out the bonus using a threshold that is a weighted sum of the sales quantity and inventory level. Specifically, the agent receives a bonus if the realized sales quantity s and inventory level i satisfy $s + \kappa_1 \cdot i \geq \kappa_2$, for some $\kappa_1, \kappa_2 \geq 0$.

(b) *Weighted-difference threshold compensation plan*: This compensation plan is similar to a weighted-sum threshold compensation plan, except that the threshold for bonus payout is a weighted *difference* of the sales quantity and inventory level. Specifically, the agent receives a bonus if the sales quantity s and inventory level i satisfy $s - \kappa_1 \cdot i \geq \kappa_2$, for some $\kappa_1, \kappa_2 \geq 0$.

Both of the above contracts are monotone and simple to implement. The former capture a general "substitutable" structure between inventory and sales, captured by the downward-sloping nature of the curve. Weighted-difference threshold compensation plans capture a "complementary" structure between the signals.

Another direction for approximating the mast-and-sail contract is to approximate the downward slope of the sail part of the optimal contract structure. We consider two possibilities:

(a) *Mast-and-flat-sail compensation plan*: This compensation plan has both a "mast" and a "sail" except that the bottom of the sail is flattened. It has two parameters, ξ_m and ξ_s, where $0 < \xi_m \le \xi_s < 1$, such that the agent receives a bonus under one of the following two conditions: (a) the sales quantity $s \ge \xi_m$ and all inventory is cleared, and (b) the sales quantity $s \ge \xi_s$ and not all inventory is cleared.

(b) *Mast-and-linearized-sail compensation plan*: This simple contract most closely mimics the structure of the mast-and-sail contract and has three parameters, ξ_m, κ_1 and κ_2, all nonnegative, such that the agent receives a bonus under one of the following two conditions: (a) the sales quantity $s \ge \xi_m$ and all inventory is cleared or (b) the sales quantity s and realized inventory level i satisfy $s + \kappa_1 \cdot i \ge \kappa_2$ and not all inventory is cleared.

One can view the mast-and-flat-sail compensation plan as a "demand-censoring aware" threshold policy because it can be viewed as a generalized sales-quota-bonus compensation plan with two thresholds, depending on whether inventory is cleared. Similarly, the mast-and-linearized sail compensation plan is a "demand-censoring aware" weighted sum compensation plan. In this case, however, the linear portion is always downward sloping, because it mimics the decreasing function s^*.

We evaluate the performance of the above three compensation plans through extensive numerical experiments, and illustrate two representative scenarios in Figure 11. We draw the following observations. First, the performance of the weighted-sum threshold compensation plan is similar to that of the corner (or sales-quota-bonus) compensation plan in that it is near optimal when the marketing and operational activities are sufficiently complementary (see Figure 11(a)), yet far from the optimal otherwise (see Figure 11(b)). Interestingly, the weighted-difference threshold compensation plan performs reasonably well in the latter case.

Second, the mast-and-flat-sail compensation plan, despite having both a "mast" and a "sail," can be outperformed by a "triangular sail" of either the weighted-sum threshold or weighted-difference threshold type (see Figure 11(a)). However, the shape of the triangular sail is important here. Contracts with two "pieces" (a mast and something resembling a sail) can tailor incentives to the "cleared inventory" and "not cleared inventory" cases in isolation from each other. A single "piece" contract (like a weighted threshold compensation plan) needs to effectively "bridge" between these

two cases by designing the transition from "cleared inventory" to "not cleared inventory." A perfect "bridge" is not always possible to build, but our experiments show single-piece contracts with strong performance are possible if constructed correctly. The correct construction (whether a weighted-sum or weighted-difference threshold) is the one that most closely mimics the structure of the underlying optimal mast-and-sail contract. For instance, a weighted-difference contract mimics the case in which the bottom tip of the mast is lower than the bottom tip of the sail.

This leads to our final observation. The performance of the mast-and-linearized-sail compensation plan is consistently near optimal (or optimal). It treats the "cleared inventory" case and "not cleared inventory" case individually and with relatively little restriction (except that the sail is triangular). Thus, it gains many of the benefits of the mast-and-sail compensation plan but remains simple and easy to compute.
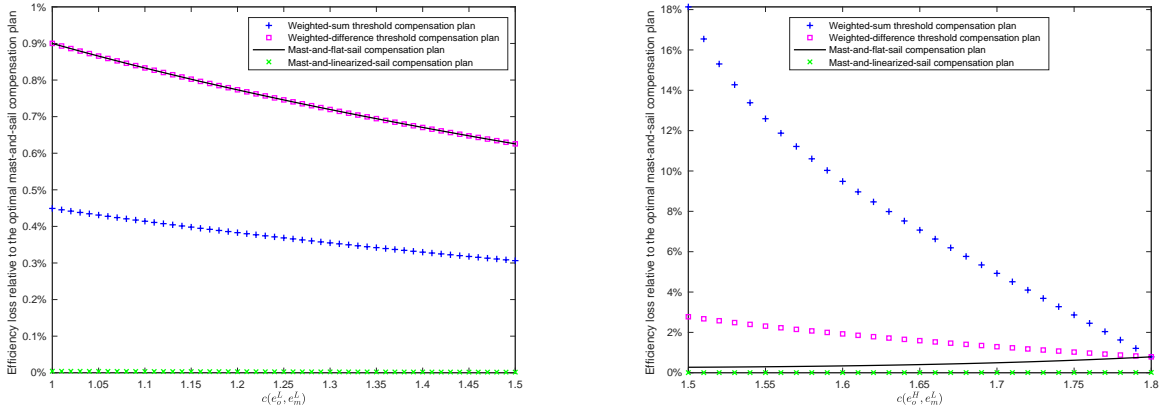
(a) Parameters: $\bar{w} = 10$, $c(e_o^H, e_m^L) = c(e_o^L, e_m^H) = 3.0$, and $c(e_o^H, e_m^H) = 3.5$. We vary $c(e_o^L, e_m^L)$ between 1 and 1.5.

(b) Parameters: $\bar{w} = 10$, $c(e_o^L, e_m^L) = 1$, $c(e_o^L, e_m^H) = 1.8$, and $c(e_o^H, e_m^H) = 3.5$. We vary $c(e_o^H, e_m^L)$ between 1.5 and 1.8.

Figure 11 **Performance of the optimal weighted-sum threshold , weighted-difference threshold, mast-and-flat-sail, and mast-and-linearized-sail compensation plans, relative to the optimal mast-and-sail compensation plan. We assume the same random distributions for I and S as in Figure 9.**

8. Endogenizing the Compensation Ceiling and Initial Inventory

In our base model, for cleanliness of analysis, we assumed $\bar{I}$ and $\bar{w}$ are exogenous. Although these assumptions are fairly standard in the literature, we relax them in this section for two purposes. One is to show the flexibility of our analytical framework. Indeed, we show the core ideas in Section 4 can be applied broadly and offer a clarifying perspective even as more complexity is introduced. Second, we derive additional managerial insights into how the firm's choice of inventory and compensation structure interplay with the underlying incentive issue.

8.1. The Role of $\bar{w}$ and More General Resource Constraints

The role of the upper bound $\bar{w}$ on compensation is a delicate one in our model. As mentioned in Section 3, the assumption is not uncommon in the literature and has been justified elsewhere. However, because $\bar{w}$ is exogenous to the model, a question remains of how to interpret it. For instance, can the firm set $\bar{w}$? If so, how high or low should it be set? How does $\bar{w}$ change the structure of the optimal compensation plan?

A few things are clear. Changing $\bar{w}$ does not change the optimality of the mast-and-sail compensation plan, because those results are parametric in $\bar{w}$. Changing $\bar{w}$ does, however, change the relative length of the mast and shape of the sail and the probability of the agent receiving the bonus under the target actions.[13]

If we view $\bar{w}$ purely as a choice of the firm, and consider its optimization over the choice of $\bar{w}$, it is obvious that larger choices of $\bar{w}$ are better. Indeed, $\bar{w}$ only enters in the constraint $w(i,s) \leq \bar{w}$, and so increasing $\bar{w}$ can only improve the objective value of the firm. This is a slippery slope. If the choice of $\bar{w}$ is unconstrained, it will be sent to infinity. When $\bar{w} = +\infty$, an optimal compensation plan need not exist in general. Finding a natural upper bound for $\bar{w}$, then, will be important when discussing its choice.

In this direction, a first step is to consider a more general constraint for bounding compensation. Let $m(i,s)$ be an L^1 function that represents the available resources for compensation by the firm when outcome (i,s) prevails; that is, constraint $w(i,s) \leq \bar{w}$ is replaced by constraint $w(i,s) \leq m(i,s)$ for almost all (i,s). For example, $m(i,s) = \alpha rs$ is an α proportion of the revenue rs when s units are sold at sales price r that can be granted to the agent as a "bonus" when targets are met.

We now show our model can be adjusted to the setting with resource constraint $w(i,s) \leq m(i,s)$. Define a new variable $\beta(i,s)$ where $w(i,s) = \beta(i,s)m(i,s)$ and $\beta(i,s) \in [0,1]$ for almost all (i,s). The new function β can be interpreted as the percentage of the resource that is given to the store manager as a bonus. Under this reformation, the problem becomes

$$\max_{\beta} \quad r \int_i \int_s sf(i|e_o^*)g(s|e_m^*)\,di\,ds - \int_i \int_s \beta(i,s)m(i,s)f(i|e_o^*)g(s|e_m^*)\,di\,ds \tag{24a}$$

$$\text{st} \quad \int_i \int_s \beta(i,s)m(i,s)f(i|e_o^*)g(s|e_m^*)\,di\,ds - c(e_o^*,e_m^*) \geq \underline{U} \tag{24b}$$

$$\int_i \int_s \beta(i,s)m(i,s)f(i|e_o^*)g(s|e_m^*)\,di\,ds - \int_i \int_s \beta(i,s)m(i,s)f(i|e_o)g(s|e_m)\,di\,ds \tag{24c}$$
$$\geq c(e_o^*,e_m^*) - c(e_o,e_m) \text{ for all } (e_o,e_m)$$

$$0 \leq \beta(i,s) \leq 1 \text{ for all } (i,s). \tag{24d}$$

---

[13] Tracking the relationship between $\bar{w}$ and the probability of bonus payout is used in Section OA.7 to provide a microfoundation for $\bar{w}$. We pursue a different microfoundation for $\bar{w}$ later in Section 8.1.1.

This problem is essentially of the form (4), except now interpreting $f(\bar{x}|\bar{a}^*)$ in that formulation as $m(i,s)f(i|e_o)g(s|e_m)$. Note the latter may not be a probability density function; however, this is irrelevant to the underlying theory, which is based on L^∞ optimization in Section IV.12 of Barvinok (2002). It suffices for $m(i,s)$ to be an L^1 function (as we assume above).[14] Thus, an optimal bang-bang contract exists for (24) with a similarly nice structure.

THEOREM 4. *There exist nonnegative multipliers ω_i and a "target" t such that an optimal solution to (24) of the following form exists:*[15]

$$w^*(i,s) = \begin{cases} m(i,s) & \text{if } \sum_{e_o,e_m} \omega_{e_o,e_m} R_{e_o,e_m}(i,s) \geq t. \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

where now

$$R_{e_o,e_m}(i,s) = 1 - \frac{\mathbb{I}[i>s]f(i|e_o)g(s|e_m) + \delta(i=s)f(i|e_o)(1-G(i|e_m))}{\mathbb{I}[i>s]f(i|e_o^*)g(s|e_m^*) + \delta(i=s)f(i|e_o^*)(1-G(i|e_m^*))}.$$

Note we do not restrict the specific form of the resource constraint. A revenue sharing constraint $(m(i,s)=\alpha is)$, for example, is a natural choice. Indeed, the size of α can explain a variety of different bonus magnitudes that we may observe in practice.

8.1.1.   Agent Incentive to "Hide" Excess Inventory. We now consider another natural choice that has the added benefit of further exploring the theme of nonmonotonicity initially discussed in Section 6.1. The nonmonotonicity result (Proposition 4 in Section 6.1) was written under the assumption that the store manager could not "hide" inventory on purpose. As discussed in that section, if the agent could do so, it brings into question the mast portion of the compensation plan for inventory realizations less than i_s. We now provide a partial justification for why that portion of the mast might prevail in practice.

Suppose we do allow the agent to hide excess inventory when they do not get the bonus but have realizations (i,s) where $i>i_m$ but $s<i$, which is precisely the case in which the agent has an incentive to hide inventory. Although we allow "hiding," we also add a penalty $\phi(i-s)$ for being found out when hiding $i-s$ units of inventory (precisely the amount to yield a bonus). Attempts to hide are not always discovered, but the larger the amount hidden, the more conspicuous, and so the probability $\rho(i-s)$ of being caught depends on $i-s$. We assume both ϕ and ρ are increasing functions with $\phi(0)=\rho(0)=0$ with a discontinuity of 0, where ϕ and ρ are right-continuous.

The bonus $\overline{w}$ can be set to discourage hiding by setting

$$-\phi(i-s)\rho(i-s) + \bar{w}(1-\rho(i-s)) \leq \underline{w} = 0$$

---

[14] In general, the product of two L^1 functions needs not be in L^1; however, because $f(i|e_o)g(s|e_m)$ is bounded and the product of an L^1 function with a bounded L^1 function is in L^1, no further restriction of m is needed.

[15] Again, this assumes the function $\sum_{i=1}^m \omega_i R_i(\vec{x})$ has zero mass at the cutoff t.

when $i = s$; that is,

$$\bar{w} \leq \frac{\phi(i-s)\rho(i-s)}{(1-\rho(i-s))}. \tag{26}$$

As discussed above, the firm has an incentive to increase $\bar{w}$ as high as possible, and so the upper-bound function $m(i,s) = \frac{w+\phi(i-s)\rho(i-s)}{(1-\rho(i-s))}$ will be chosen and suffices to prevent the hiding of inventory. When $i = s$, set

$$m(i,s) = \lim_{s \to i^+} \frac{\phi(i-s)\rho(i-s)}{1-\rho(i-s)} = \bar{w},$$

which is well-defined because of the right continuity of ϕ and ρ at zero. This provides another interpretable choice for the resource constraint $m(i,s)$ that also bounds (and effectively endogenizes) the choice of $\bar{w}$.

Also observe that (26) can explain how, in practice, bonuses may be quite small. If the probability of getting caught is quite small, a large bonus is never forthcoming from (26), because a large bonus gives a high incentive to "cheat," and cheating cannot easily be caught. In a large decentralized chain of stores where oversight may be difficult, we expect very modest bonuses to be the norm. On the other hand, in settings where the store manager is very likely to be caught if hiding inventory, larger bonuses prevail.

8.2. Endogenizing Initial Inventory

In this section, we endogenize the choice of initial inventory $\bar{I}$. Whether being under the control of the firm or the store manager is more natural for $\bar{I}$ is a matter of debate. In this paper, we analyze the former. This perspective particularly applies in settings where the firm oversees a large chain of small stores where forecasting and ordering are done centrally. If the choice of inventory is given to the store manager, the potential exists for additional incentive issues that go beyond our scope. Moreover, this focus allows us to compare the choice of inventory with that of the classical newsvendor, assuming a first-best outcome for the incentive issue.

To set the benchmark, suppose we can ignore the incentive compatibility constraint of the store manager and pay her a constant wage to meet the minimum utility $\underline{U}$ to work at effort level (e_o^H, e_m^H). Under this assumption, the firm's problem is

$$\max_{\bar{I}} r\mathbf{E}[S|\bar{I}] - C(\bar{I}), \tag{27}$$

where $C(\cdot)$ is a convex increasing cost for procuring inventory $\bar{I}$. Thus, the optimal inventory level is the classical newsvendor solution I^{NV} that solves

$$r\frac{d}{d\bar{I}}\mathbf{E}[S|I^{\text{NV}}] = C'(I^{\text{NV}}). \tag{28}$$

As for the second-best compensation plan, where the store manager's incentives must be taken into consideration, the firm's inventory decision becomes

$$\max_{\bar{I}} r\mathbf{E}[S|\bar{I}] - W(\bar{I}) - C(\bar{I}), \tag{29}$$

where

$$W(\bar{I}) \triangleq \min_{w} \quad \mathbb{E}[w(I,S)|e_o^H, e_m^H]$$
$$\text{s.t.} \quad \mathbb{E}[w(I,S)|e_o^H, e_m^H] - c(e_o^H, e_m^H) \geq \underline{U}$$
$$\mathbb{E}[w(I,S)|e_o^H, e_m^H] - \mathbb{E}[w(I,S)|e_o, e_m] \geq c(e_o^H, e_m^H) - c(e_o, e_m) \text{ for all } (e_o, e_m)$$
$$0 \leq w(i,s) \leq \bar{w} \text{ for all } (i,s)$$

is the optimal value function of the expected wage payout problem when (e_o^H, e_m^H) is the target effort level.

PROPOSITION 7. *(i) Under the assumption that (e_o^H, e_m^H) is the target effort level and f and g satisfy the MLRP, we have $\frac{d}{d\bar{I}}W(\bar{I}) < 0$ for all $\bar{I}$. That is, an increase in $\bar{I}$ leads to a decrease in the expected payout to the store manager.*

(ii) The firm's optimal inventory level, by accounting for the multitasking store manager problem, is higher than that in the newsvendor problem, which helps the firm achieve a lower expected payment to the store manager than otherwise.

The intuition behind Proposition 7(i) is as follows. Because the firm is more likely to pay a bonus if the inventory is cleared (as long as sales are greater than i_m), increasing inventory reduces the chance of inventory clearing and thus the chance of paying out the bonus.

Proposition 7(ii) entails optimizing the initial inventory level $\bar{I}$. Let I^* be the optimal choice of inventory in (29). The first-order condition of (29) yields the necessary optimality condition for I^*:

$$r\frac{d}{d\bar{I}}\mathbf{E}[S|I^*] - \frac{d}{d\bar{I}}W(I^*) = C'(I^*). \tag{30}$$

In light of (29) and (30), and because $\frac{d}{d\bar{I}}W(I^*) < 0$ (by Proposition 7) and C is a convex increasing function, we can conclude $I^* > I^{NV}$. In other words, the firm over-invests in inventory as compared to the traditional newsvendor setting when the effort level of the store manager cannot be contracted on. A higher inventory level benefits the firm by reducing the complications in contract design due to the information distortions that arise from demand censoring, because higher levels of inventory mitigate the possibility of censoring.

9.   Concluding Remarks

We have developed a theory of incentives under multitasking, mirroring the problem of compensating a store manager who divides her effort between maintaining inventory (operations) and generating demand (marketing), a topic of well-recognized importance that to date has been primarily studied empirically. Our formulation of this *multitasking store manager problem* reflects several of its fundamental features: the agent has to simultaneously choose and exert multiple types of effort that stochastically influences the outcome, which is the "weaker link" between inventory and demand outcomes. When inventory is the weaker link, *demand censoring* arises, further crippling the firm's access to meaningful information about the agent's efforts.

In our attempt to overcome the technical challenges associated with this multitasking store manager problem, we developed a bang-bang control approach to contract theory. We believe this approach has the potential to apply to myriad incentive-design problems. Specific to the multitasking store manager problem, we derive an optimal compensation plan that exhibits a "mast-and-sail" shape and provides rich managerial interpretations along several dimensions, including the nonmonotonicity of the compensation plan, the relationship between operational and marketing outcomes, and the non-optimality of two simpler compensation plans. We also discuss ways to approximate the optimal mast-and-sail compensation plan that may entail either omitting the mast or linearizing the sail, generating practical implications for the incentive-design problem.

We conclude our paper by mentioning a few interesting extensions that are beyond the scope of the current analysis. So far, consistent with the literature on moral hazard with limited liability, we have assumed both the firm and store manager are risk neutral. When the agent is risk averse, other methodologies must be considered; see Ke and Ryan (2018) and Kirkegaard (2017a, 2017b) for example. On the other hand, a number of researchers (e.g., Innes 1990; Oyer 2000; Poblete and Spulber 2012) study the *single-tasking* setting by adding an *a priori* restriction that the contract be monotone. This analysis has a different flavor than our approach. We should point out that, to our knowledge, deriving the structure of an optimal monotone contract in the multitasking setting is not known in general. We leave it to future work to explore this direction.

Appendix: Proofs

Proof of Lemma 1. Due to constraint (4d), we may assume w is a function in $L^\infty(\mathcal{X})$, the space of uniformly bounded functions on $\mathcal{X}$. Moreover, those same constraints ensure the feasible region of (4) is bounded in the norm on $L^\infty(\mathcal{X})$. Hence, by Alaoglu's Theorem (see Theorem 5.105 in Aliprantis and Border 2006), the set $\{w : 0 \le w(\vec{x}) \le \overline{w}\}$ is compact in the weak topology $\sigma(L^\infty(\mathcal{X}), L^1(\mathcal{X}))$ (for a definition of this weak topology, see Section 5.14 of Aliprantis and Border (2006)). Because $f(\cdot | \vec{a})$ is in $L^1(\mathcal{X})$ for all $\vec{a} \in \mathcal{A}$, the constraints (4b) and (4c) are continuous in the $\sigma(L^\infty(\mathcal{X}), L^1(\mathcal{X}))$ topology, and so the feasible region is a closed subset of $\{w : 0 \le w(\vec{x}) \le \overline{w}\}$ and thus also compact in the $\sigma(L^\infty(\mathcal{X}), L^1(\mathcal{X}))$ topology. Moreover, the objective function $V(w, a)$ is continuous in the $\sigma(L^\infty(\mathcal{X}), L^1(\mathcal{X}))$ topology, and so, by Weierstrass's Theorem (see Theorem 2.35 in Aliprantis and Border 2006), an optimal contract exists. □

Proof of Lemma 2. This argument can be adapted from Proposition III.5.3 in Barvinok (2002). Our problem (4) is a linear program in $L^\infty(\mathcal{X})$ with finitely many constraints. Proposition III.5.3 in Barvinok (2002) shows that in linear programs in $L^\infty[0,1]$, extremal solutions have a bang-bang structure. This result can be adjusted to the multidimensional setting over the compact set $\mathcal{X}$ using standard arguments. Details are omitted. □

Proof of Theorem 1. Bauer's Maximum Principle (see Theorem 7.69 in Aliprantis and Border (2006)) states that every lower semicontinuous concave function has an extreme-point minimizer over a compact convex set. The feasible region W is convex because all constraints are linear. The compactness of W and continuity of the objective of (4) were argued in the proof of Lemma 1. Hence, by Bauer's Maximum Principle, an optimal extremal contract exists. Therefore, by Lemma 2, an optimal bang-bang contract exists. □

Proof of Theorem 2. We use Barvinok (2002) Proposition IV.12.6 — which is based on duality and complementary slackness — to characterize the structure of an optimal solution to our linear program over $L^\infty[0,1]$. Adapted to our setting, this method involves setting a Lagrange multiplier μ for the IR constraint (4b) and Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ for the constraints in (4c). Given a choice of nonnegative dual multipliers μ and λ_i we define a function

$$p(\vec{x}) \triangleq \max\{0, -f(\vec{x}|\vec{a}^*) + \mu f(\vec{x}|\vec{a}^*) + \sum_{i=1}^{m} \lambda_i R_i(\vec{x}) f(\vec{x}|\vec{a}^*)\} \tag{31}$$

and, by Barvinok (2002, Proposition IV.12.6), an optimal bang-bang contract has the form

$$w^*(\vec{x}) = \begin{cases} \overline{w} & \text{if } p(\vec{x}) > 0 \\ 0 & \text{if } p(\vec{x}) = 0 \end{cases}, \tag{32}$$

assuming $\{\vec{x} : -f(\vec{x}|\vec{a}^*) + \mu f(\vec{x}|\vec{a}^*) + \sum_{i=1}^{m} \lambda_i R_i(\vec{x}) f(\vec{x}|\vec{a}^*) = 0\}$ has measure zero.

Observe that $p(\vec{x}) > 0$ if and only if

$$-1 + \mu + \sum_{i=1}^{m} \lambda_i R_i(\vec{x}) > 0, \tag{33}$$

which can be rewritten as

$$\sum_{i=1}^{m} \omega_i R_i(\vec{x}) > \tfrac{1-\mu}{\sum_{i=1}^{m} \lambda_i},$$

where $\omega_i \triangleq \frac{\lambda_i}{\sum_{i=1}^m \lambda_i}$. In the first step, we divide both sides of the argument in the "max" in (31) by $f(\vec{x}|\vec{a}^*)$, which is positive for all $\vec{x} \in \mathcal{X}$. In the last step, we divide through by $\sum_{i=1}^m \lambda_i$ which we assume is nonzero. This assumption is without loss, because otherwise $\lambda_i = 0$ for all i, and so (33) either holds for all x or no x. Thus, if $\sum_{i=1}^m \lambda_i = 0$, the extremal contract is either constant at 0 or constant at $\bar{w}$. In either case, the contract is either not feasible (violates IR) or not optimal (pays out the maximum) and thus can be excluded from consideration. Observe that ω_i is nonnegative because all the λ_i are nonnegative. Using this equivalence and defining

$$t \triangleq \frac{1-\mu}{\sum_{i=1}^m \lambda_i},$$

we may re-express the optimal bang-bang contract in (32) as[16]

$$w^*(\vec{x}) = \begin{cases} \overline{w} & \text{if } \sum_{i=1}^m \omega_i R_i(\vec{x}) \geq t. \\ 0 & \text{otherwise.} \end{cases}$$

This completes the proof. □

Proof of Theorem 3. By Theorem 2, an optimal contract to (4) exists that is feasible to (8). Thus, the optimal value of (8) is at least the optimal value of (4). Moreover, because the feasible region of (8) is a restriction of the feasible region of (4) (i.e., it restricts to information-trigger contracts), the value of the former cannot exceed the value of latter. Together, this implies both problems have the same optimal value. An optimal solution (ω^*, t^*) of (8) yields the trigger contract

$$w^*(\vec{x}) = \begin{cases} \overline{w} & \text{if } \sum_{i=1}^m \omega_i^* R_i(\vec{x}) \geq t^* \\ 0 & \text{otherwise,} \end{cases}$$

which is a feasible solution to (4). Moreover, w^* attains the optimal value of (4) because (ω^*, t^*) is optimal to (8) and the values of both problems are equal. Thus, w^* is an optimal contract to (4). □

Proof of Lemma 3. When the context is clear, we will lighten notation to $\Pr(I \leq i, S \leq s)$. First, note that

$$\Pr(I \leq i, S \leq s) = \Pr(I \leq i, \min\{Q, I\} \leq s) = \Pr(I \leq i, Q < I, Q \leq s) + \Pr(I \leq s, Q \geq I)$$

because $s \leq i$. We can develop this derivation further by noting that

$$\begin{aligned} \Pr(I \leq i, Q < I, Q \leq s) &= \Pr(Q < I \leq i, Q \leq s) \\ &= \int_0^s \int_q^i f(j|e_o) g(q|e_m) \, dj \, dq \\ &= \int_0^s (F(i|e_o) - F(q|e_o)) g(q|e_m) \, dq \\ &= F(i|e_o) \int_o^s g(q|e_m) \, dq - \int_0^s F(q|e_o) g(q|e_m) \, dq \\ &= F(i|e_o) G(s|e_m) - \int_0^s F(q|e_o) g(q|e_m) \, dq. \end{aligned}$$

It is useful to further analyze this by integration by parts to conclude that

$$\Pr(I \leq i, Q < I, Q \leq s) = F(i|e_o) G(s|e_m) - F(s|e_o) G(s|e_m) + \int_0^s G(q|e_m) f(q|e_o) \, dq. \tag{34}$$

---

[16] Note we change the strict inequality in (32) to a weak inequality here. Because $p(\vec{x}) = 0$ is assumed to be a measure-zero event, this change can be made without loss.

Moreover,

$$
\begin{aligned}
\Pr(I \leq s, Q \geq I) &= \int_0^s \left(\int_j^{\bar{Q}} g(q|e_m) dq \right) f(j|e_o) dj \\
&= \int_0^s \left(1 - G(j|e_m) \right) f(j|e_o) dj \\
&= F(s|e_o) - \int_0^s G(j|e_m) f(j|e_o) dj.
\end{aligned}
\tag{35}
$$

From (34) and (35), we can conclude that

$$
\begin{aligned}
\Pr(I \leq i, S \leq s) &= \Pr(I \leq i, Q < I, Q \leq s) + \Pr(I \leq s, Q \geq I) \\
&= F(i|e_o) G(s|e_m) - F(s|e_o) G(s|e_m) + \int_0^s G(q|e_m) f(q|e_o) dq + F(s|e_o) - \int_0^s G(j|e_m) f(j|e_o) dj \\
&= F(s|e_o) + G(s|e_m)[F(i|e_o) - F(s|e_o)]
\end{aligned}
$$

for $s < i$, and that if $s = i$,

$$
\Pr(I \leq i, S \leq s) = F(i|e_o).
$$

To conclude, we have derived the joint cumulative distribution function as

$$
\Pr(I \leq i, S \leq s) = \begin{cases} F(s|e_o) + G(s|e_m)[F(i|e_o) - F(s|e_o)] & \text{if } s < i \\ F(i|e_o) & \text{if } s = i \end{cases}
$$

and complete the proof. $\qquad \square$

Proof of Proposition 1. We first note that the domain of s^* need not be all $[0, \bar{I}]$, because for a given i, an s might not exist such that $\varphi^{\text{NSO}}(i, s) = t$. However, properties on the $R_{e_o, e_m}^{\text{NSO}}$ imply that once some $\bar{i}$ and s exist such that $\varphi(\bar{i}, s) = t$ is nonempty, the same is true for any i larger than $\bar{i}$. That is, the domain of s^* is an interval of the form $[\bar{i}, \bar{I}]$.

Next, we show that the mapping $s^*(i)$ is well-behaved. Specifically, it is a decreasing, continuous, and almost everywhere differentiable function of i on its domain. The reasoning is as follows. As described in the paragraph above (20), each of the $R_{e_o, e_m}^{\text{NSO}}(i, s)$ are continuous, increasing, and nonconstant in each of its coordinates. Hence, the same is true of the function $\varphi(i, s)$. Hence the level set $\{(i, s) : \varphi(i, s) = t\}$ has the structure illustrated in Figure 3. That is, the level set $\{(i, s) : \varphi(i, s) = t\}$ is the region between two decreasing and continuous functions. Observe that the graph of $s^*(i)$ in (i, s)-space is precisely the lower envelope of $\{(i, s) : \varphi(i, s) = t\}$. Thus, we can conclude that s^* is a decreasing, continuous, and almost everywhere differentiable function of i.

Finally, we prove the existence of uniqueness of i_s. Because $s^*(i)$ is a decreasing and continuous function of i and we have assumed B^{NSO} has positive measure, the $\arg\min$ in (22) is nonempty (by Brouwer's Fixed Point Theorem (Corollary 17.56 in Aliprantis and Border (2006)) and is a singleton (because s^* is decreasing and the 45° line is strictly increasing). Hence, a unique choice exists for i_s. $\qquad \square$

Proof of Proposition 2. To see that the set of i such that $\varphi^{\text{SO}}(i) = t$ is nonempty, observe that $\varphi^{\text{SO}}(\bar{I}) \geq t$ because $(\bar{I}, \bar{I})$ must be in every bonus region, and an i exists such that $\varphi^{\text{SO}}(i) < 0$. The latter follows because properties (17) and (18) imply $\int_{D^{\text{SO}}} R_{e_o, e_m}^{\text{SO}}(i, s) f(i|e_o) g(i|e_m) ds di \leq 0$ for all e_o, e_m, and so for all e_o, e_m an

i exists such that $R^{\text{SO}}_{e_o,e_m}(i) < 0$. Hence, by this continuity of $\varphi^{\text{SO}}(i)$, the set of i such that $\varphi^{\text{SO}}(i) = t$ is nonempty. Again, from the monotonicity properties of φ^{SO}, the bonus region in the stockout case is precisely as defined in (23). $\qquad\square$

Proof of Proposition 3. From the definition of $s^*(i)$, i_s, and i_m, the following holds for every valid choice of t and $(\omega_{e_o^H,e_m^L}, \omega_{e_o^L,e_m^L})$:

$$\omega_{e_o^H,e_m^L} \frac{1-G(i_m|e_m^L)}{1-G(i_m|e_m^H)} + \omega_{e_o^L,e_m^H} \frac{f(i_m|e_o^L)}{f(i_m|e_o^H)} + \omega_{e_o^L,e_m^L} \frac{f(i_m|e_o^L)(1-G(i_m|e_m^L))}{f(i_m|e_o^H)(1-G(i_m|e_m^H))}$$

$$= \omega_{e_o^H,e_m^L} \frac{g(s^*(i_s)|e_m^L)}{g(s^*(i_s)|e_m^H)} + \omega_{e_o^L,e_m^H} \frac{f(i_s|e_o^L)}{f(i_s|e_o^H)}) + \omega_{e_o^L,e_m^L} \frac{f(i_s|e_o^L)g(s^*(i_s)|e_m^L)}{f(i_s|e_o^H)g(s^*(i_s)|e_m^H)}$$

$$= \omega_{e_o^H,e_m^L} \frac{g(i_s|e_m^L)}{g(i_s|e_m^H)} + \omega_{e_o^L,e_m^H} \frac{f(i_s|e_o^L)}{f(i_s|e_o^H)}) + \omega_{e_o^L,e_m^L} \frac{f(i_s|e_o^L)g(i_s|e_m^L)}{f(i_s|e_o^H)g(i_s|e_m^H)},$$

where the second equality follows because $s^*(i_s) = i_s$. MRLP distributions also have increasing failure rates, so we have $\frac{1-G(\cdot|e_m^L)}{1-G(\cdot|e_m^H)} < \frac{g(\cdot|e_m^L)}{g(\cdot|e_m^*)}$, which implies

$$\begin{aligned}
&\omega_{e_o^H,e_m^L} \frac{g(i_s|e_m^L)}{g(i_s|e_m^H)} + \omega_{e_o^L,e_m^H} \frac{f(i_s|e_o^L)}{f(i_s|e_o^H)}) + \omega_{e_o^L,e_m^L} \frac{f(i_s|e_o^L)g(i_s|e_m^L)}{f(i_s|e_o^H)g(i_s|e_m^H)} \\
&\leq \omega_{e_o^H,e_m^L} \frac{g(i_m|e_m^L)}{g(i_m|e_m^H)} + \omega_{e_o^L,e_m^H} \frac{f(i_m|e_o^L)}{f(i_m|e_o^H)}) + \omega_{e_o^L,e_m^L} \frac{f(i_m|e_o^L)g(i_m|e_m^L)}{f(i_m|e_o^H)g(i_m|e_m^H)}.
\end{aligned} \tag{36}$$

The inequality is strict if $\min\{\omega_{e_o^H,e_m^L}, \omega_{e_o^L,e_m^L}\} > 0$. Because f and g satisfy the MLRP, $\omega_{e_m^H,e_m^L} \frac{g(i|e_m^L)}{g(i|e_m^H)} + \omega_{e_o^L,e_m^H} \frac{f(i|e_o^L)}{f(i|e_o^H)}) + \omega_{e_o^L,e_m^L} \frac{f(i|e_o^L)g(i|e_m^L)}{f(i|e_o^H)g(i|e_m^H)}$ is a decreasing function of i. As result, from (36), we can conclude $i_1^* \geq i_2^*$ and $i_1^* > i_2^*$ if $\min\{\omega_{e_o^H,e_m^L}, \omega_{e_o^L,e_m^L}\} > 0$. $\qquad\square$

Proof of Proposition 4. We first prove (i). For $i < i_m$, notice that $w^*(s',i) = w^*(s'',i) = 0$ for all $(i,s'),(i,s'') \in D$ with $s' \leq s''$. For $i \in [i_m, i_s]$, observe that $w(s,i) = 0$ if $s < i$ and $\bar{w}$ if $s = i$. Thus, $w^*(s',i) \leq w^*(s'',i) = 0$ for all $(i,s'),(i,s'') \in D$ with $s' \leq s''$. Finally, suppose $i > i_s$. By the definition of B^{NSO} and B^{SO} in (19) and (23) , $w^*(s,i) = \bar{w}$ if $s^*(i) \leq s \leq i$, and 0 otherwise. Hence, again, $w^*(s',i) \leq w^*(s'',i) = 0$ for all $(i,s'),(i,s'') \in D$ with $s' \leq s''$. This implies w^* is monotone in s.

We now show (ii). First, we show w^* is monotone in i. Suppose $s \geq i_s$. This implies $w(i,s) = \bar{w}$ for all $i \geq s$ (which is needed for $(i,s) \in D$) and so w^* is monotone in i in this region. Suppose otherwise that $s < i_s$. In this case, $w(i,s) = 0$ for $i \leq \min(s^*)^{-1}(s)$ and $\bar{w}$ otherwise, which again yields monotonicity in i. Joint monotonicity now follows from monotonicity in both direction i and s.

Turning now to (iii), if $i_m < i_s$, an $\epsilon > 0$ exists such that $w^*(i_m, i_m) = \bar{w}$ but $w^*(i_m + \epsilon, i_m) = 0$, and thus w^* is not monotone in i.

Finally, we show w^* also fails joint monotonicity when $i_m < i_s$. Consider the point $(i^\circ, s^\circ) \triangleq (i_s + \epsilon, s^*(i_1 + \epsilon))$ on the graph of s^*. This means $w^*(i^\circ, s^\circ) = \bar{w}$ because the graph of s^* for $i \geq i_s$ lies in the bonus region of w^*. Choose any (i,s) in the open line segment between (i_m, i_m) and (i°, s°) with $i < i_s$. Such a choice is possible because $i_m < i_s$. See Figure 6 for an illustration.

Note that (i,s) does not lie in the bonus region. Indeed, $i < i_s$ (by construction) and $s^*(i) < i_s$ (because $s^\circ \leq i_s$ because s^* is a decreasing function). Hence, $w^*(i,s) = 0$. This yields a contradiction of monotonicity. Observe that $(i_m, i_m) < (i,s) < (i^\circ, s^\circ)$ but $w^*(i_m, i_m) = w^*(i^\circ, s^\circ) = \bar{w}$ and $w^*(i,s) = 0$.[17] $\qquad\square$

---

[17] Not this proof allows for the possibility that $s^*(i) = s^*(i_s)$ for all $i \geq i_s$, which cannot be ruled out under the MLRP assumptions.
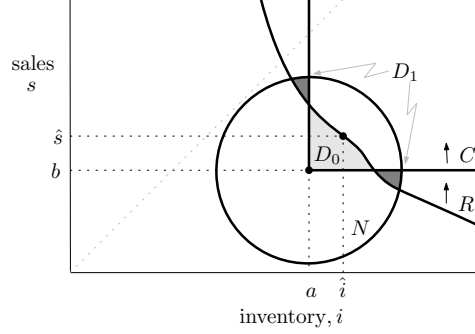
Figure 12   **Illustration of the objects in the proof of Proposition 6.**

Proof of Proposition 6. We prove both by noting that no optimal compensation plan can have a bonus region that has a "corner," defined as follows. Let B denote the bonus region of an optimal contact, that is, where the store manager receives a positive bonus. Let (a,b) be a point in B and let $C \triangleq (a,b) + \mathbb{R}^n_+$ denote the (translated) cone of points that are pointwise no smaller than $(\bar{i}, \bar{s})$. We say (a,b) is a corner point of B if (a,b) is an isolated extreme point of B. That is, a nonempty neighborhood N of $(\bar{i}, \bar{s})$ exists, such that $B \cap N = C \cap N$. This claim rules out corner compensation plans, but also any compensation plan with a "corner" as defined above.

We first prove the claim in the simplest case, in which on the no-jump constraint for (e_o^L, e_m^L) is tight, the other two no-jump constraints (for (e_o^H, e_m^L) and (e_o^L, e_m^H)) are slack. This case is the easiest to understand and will make clear what needs to be handled in the more challenging settings. Let w^* be an optimal compensation plan with a corner at (a,b) go to a new point $(\hat{i}, \hat{s}) = (a,b) + \delta(1,1)$, where δ is chosen sufficiently small so that the no-jump constraints for (e_o^H, e_m^L) and for (e_o^L, e_m^H) remain slack. Now, define the set $R = \left\{ (i,s) : R_{e_o^L, e_m^L}^{\text{NSO}}(i,s) \geq R_{e_o^L, e_m^L}^{\text{NSO}}(\hat{i}, \hat{s}) \right\}$, which is the $R_{e_o^L, e_m^L}^{\text{NSO}}(\hat{i}, \hat{s})$-superlevel sets of $R_{e_o^L, e_m^L}^{\text{NSO}}(i,s)$. Now, δ is also chosen sufficiently small so that both of the sets

$$D_0 \triangleq \{(i,s) : s < i\} \cap (C \setminus R)$$

$$D_1 \triangleq \{(i,s) : s < i\} \cap (R \cap N \setminus C)$$

have positive measure (where N is defined when we say (a,b) is a corner). We intersect both sets with $\{(i,s) : s < i\}$ so that we only deal with points in the interior of the domain of the optimal compensation plan w^* (which is only defined on $\{(i,s) : i \leq s\}$). Such a choice for δ is possible because the boundary of the set R expressed by the implicit function theorem by $r(i)$ as a function of i is a *strictly decreasing* function of i. The fact that $r(i)$ is both a function and is strictly decreasing is due to the assumption that f and g satisfy the strict MLRP. See Figure 7(a) for a visual representation of the sets D_0 and D_1.

Now, consider the perturbation function h defined as follows:

$$h(s,i) = \begin{cases} -\epsilon_0 & \text{if } (i,s) \in D_0 \\ \epsilon_1 & \text{if } (i,s) \in D_1 \\ 0 & \text{otherwise,} \end{cases}$$

where $\epsilon_0, \epsilon_1 > 0$ are chosen so that $\mathbb{E}[h|e_o^H, e_m^H] = 0$. Such a choice is possible because D_0 and D_1 both have positive measure. Now consider the new optimal compensation plan $w' = w^* + h$. We claim w' is also an optimal compensation plan. Indeed,

$$\mathbb{E}[w'|e_o^H, e_m^H] = \mathbb{E}[w^*|e_o^H, e_m^H] + \mathbb{E}[h|e_o^H, e_m^H] = \mathbb{E}[w^*|e_o^H, e_m^H]$$

because $\mathbb{E}[h|e_o^H, e_m^H] = 0$. Thus, if we can show w' is feasible, then it is optimal. Since we have assumed that the IR constraint is not binding at w^* and ϵ_0, and D_2 can be chosen sufficiently small so that the IR constraint is satisfied at w' (indeed, the condition that $\mathbb{E}[h|e_o^H, e_m^H] = 0$ is only a single linear constraint on ϵ_0 and D_2, and thus a degree of freedom from that requirement allows us to drive ϵ_0 and D_2 arbitrarily small). Now consider the IC constraint. We claim that

$$\int\limits_{s \leq i} R_{e_o^L, e_m^L}(i,s) w'(i,s) f(i|e_o^L) g(s|e_m^L) di ds > \int\limits_{s \leq i} R_{e_o^L, e_m^L}(i,s) w^*(i,s) f(i|e_o^L) g(s|e_m^L) di ds \qquad (37)$$

$$\geq c(e_o^H, e_m^H) - c(e_o^L, e_m^L)$$

holds for all (e_o, e_m). Notice the second inequality holds because w^* is feasible to the IC constraint (here we are taking the form of IC constraints from (4c)). Thus, it remains to show (37). This follows because

$$\int\limits_{s < i} R_{e_o^L, e_m^L}^{\mathrm{NSO}}(i,s) h(s,i) f(i|e_o^L) g(s|e_m^L) di ds$$

$$= -\epsilon_0 \int\limits_{D_0} R_{e_o^L, e_m^L}^{\mathrm{NSO}}(i,s) f(i|e_o^L) g(s|e_m^L) di ds + D_2 \int\limits_{D_1} R_{e_o^L, e_m^L}^{\mathrm{NSO}}(i,s) f(i|e_o^L) g(s|e_m^L) di ds$$

$$> R_{e_o^L, e_m^L}^{\mathrm{NSO}}(a,b) \left[-\epsilon_0 \int\limits_{D_0} f(i|e_o^L) g(s|e_m^L) di ds \right] + R_{e_o^L, e_m^L}^{\mathrm{NSO}}(\hat{i}, \hat{s}) \left[D_2 \int\limits_{D_1} f(i|e_o^L) g(s|e_m^L) di ds \right]$$

$$> R(a,b) \left[-\epsilon_0 \int\limits_{D_0} f(i|e_o^L) g(s|e_m^L) di ds + D_2 \int\limits_{D_1} f(i|e_o^L) g(s|e_m^L) di ds \right] = 0,$$

where the key fact in each step is that $R_{e_o^L, e_m^L}^{\mathrm{NSO}}(i,s)$ is coordinatewise strictly increasing. This implies (37). Indeed, observe that it suffices to integrate in the region $s < i$ to conclude the IC constraints bind for w' because $w'(i,s) = w^*(i,s)$ for $i = s$ (due to $h(i,s) = 0$ for $i = s$). We thus conclude w' is an optimal compensation plan.

In fact, we have shown something more in (37): the IC constraints are slack at optimal compensation plan w'. However, because problem (2) is linear, every optimal compensation plan must be on the boundary of the feasible region. This implies that either w' must bind the IR constraint, which violates our assumption of positive rents, a contradiction. This completes the proof in the special case of the argument in which only the no-jump constraint for (e_o^L, e_m^L) was initially tight.

We now allow the possibility that at least one of the other two no-jump constraints is tight at (a,b). In the case in which exactly one of the no-jump constraints for (e_o^H, e_m^L) or (e_o^L, e_m^H) is tight, an argument largely analogous to the previous one can be conducted. Here, the region D_1 will consist of only one "piece" (the D_1 in Figure 12 has two distinct pieces) because the no-jump constraints for (e_o^H, e_m^L) or (e_o^L, e_m^H) correspond to horizontal and vertical superlevel sets for $R_{e_o, e_m}^{\mathrm{NSO}}(i,s)$ from the discussion following (18). The difficulty here

is that when both of the no-jump constraints for (e_o^H, e_m^L) or (e_o^L, e_m^H) are tight, "room" remains to construct D_1 as in Figure 12. In this setting, we need to construct three regions, D_0, D_1, and D_2, where a tradeoff exists between the value of the perturbation in regions D_1 and D_2 to lead to a strictly increasing covariance as we were able to show in (37).

To make this argument, we need the following definitions. We also move a distance δ along $(1,1)$ to the point $(\hat{i}, \hat{s})$. The superlevel set R and the region D_0 are defined exactly as before. To define D_1 and D_2, we need the following concepts. At (a,b), $R_{e_o^H, e_m^L}(i,s) \triangleq R_2(s)$ is constant in i at b with value $R_1(b)$ and so we consider the horizontal line through the point (a,b). Where this line intersects R is denoted (i_1, b). The region D_1 is chosen below the horizontal line to the right of (i_1, b), above R, and inside N. Similarly, $R_{e_o^L, e_m^H}(i,s) \triangleq R_2(i)$ is constant in s at a (at the value $R_2(a)$). The region D_2 is chosen to the left of the vertical line above (a, s_2) and above the lower envelope of R. The specific sets D_1 and D_2 and δ are chosen so that an $\bar{\epsilon} > 0$ exists such that

$$\mathbb{E}[R_1(s)|D_1] - R_1(b) \leq \bar{\epsilon} \text{ and } \mathbb{E}[R_2(i)|D_2] - R_2(a) \leq \bar{\epsilon},$$

where (as used above) $\mathbb{E}[\cdot]$ is the expectation with respect to distributions with effort (e_o^H, e_m^H). From the regions D_0, D_1, and D_2, we define the perturbation

$$h(i,s) = \begin{cases} -\epsilon_0 & \text{if } (i,s) \in D_0 \\ \epsilon_1 & \text{if } (i,s) \in D_1 \\ \epsilon_2 & \text{if } (i,s) \in D_2 \\ 0 & \text{otherwise.} \end{cases}$$

Using similar logic as the simpler case above, it suffices to show that $\mathbb{E}[h(S,I)] = 0$ and $\mathbb{E}[R(S,I)h(S,I)] > 0$. We now have three degrees of freedom, and so this is possible. For the condition $\mathbb{E}[h(S,I)] = 0$, this only requires

$$\epsilon_0 \mathbb{P}(D_0) = \epsilon_1 \mathbb{P}(D_1) + \epsilon_2 \mathbb{P}(D_2),$$

where $\mathbb{P}[\cdot]$ is the probability measure with respect to distributions with effort (e_o^H, e_m^H). As for the covariance condition $\mathbb{E}[R(S,I)h(S,I)] > 0$, this analysis is more delicate. We first show how to find ϵ_1 and ϵ_2 such that

$$\frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} \left(\epsilon_1 \int_{D_1} g(s)f(i)dsdi + \epsilon_2 \int_{D_2} g(s)f(i)dsdi \right)$$
$$< \epsilon_1 \int_{D_1} R_1(s)g(s)f(i)dsdi + \epsilon_2 \int_{D_2} R_1(s)g(s)f(i)dsdi$$

and

$$\frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} \left(\epsilon_1 \int_{D_1} g(s)f(i)dsdi + \epsilon_2 \int_{D_2} g(s)f(i)dsdi \right)$$
$$< \epsilon_1 \int_{D_1} R_2(i)g(s)f(i)dsdi + \epsilon_2 \int_{D_2} R_2(i)g(s)f(i)dsdi.$$

To see that, we note the coefficient of ϵ_1 in the first inequality is

$$\int_{D_1} R_1(s)g(s)f(i)dsdi - \frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} \int_{D_1} g(s)f(i)dsdi > 0,$$

and the coefficient of ϵ_1 in the second inequality is

$$\int_{D_1} R_2(i)g(s)f(i)dsdi - \frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} \int_{D_1} g(s)f(i)dsdi < 0.$$

Therefore, the existence of such ϵ_1 and ϵ_2 depends on

$$\frac{\frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} \int_{D_2} g(s)f(i)dsdi - \int_{D_2} R_2(i)g(s)f(i)dsdi}{\int_{D_1} R_2(i)g(s)f(i)dsdi - \frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} \int_{D_1} g(s)f(i)dsdi}$$

$$> \frac{\epsilon_1}{\epsilon_2} > \frac{\frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} \int_{D_2} g(s)f(i)dsdi - \int_{D_2} R_1(s)g(s)f(i)dsdi}{\int_{D_1} R_1(s)g(s)f(i)dsdi - \frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} \int_{D_1} g(s)f(i)dsdi}.$$

The sufficient condition is

$$\frac{\frac{\int_{D_2} R_2(i)g(s)f(i)dsdi}{\int_{D_2} g(s)f(i)dsdi} - \frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi}}{\frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} - \frac{\int_{D_1} R_2(i)g(s)f(i)dsdi}{\int_{D_1} g(s)f(i)dsdi}} > \frac{\frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} - \frac{\int_{D_2} R_1(s)g(s)f(i)dsdi}{\int_{D_2} g(s)f(i)dsdi}}{\frac{\int_{D_1} R_1(s)g(s)f(i)dsdi}{\int_{D_1} g(s)f(i)dsdi} - \frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi}}.$$

We can choose D_0 such that

$$\frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} - R_2(a) \le \bar{\epsilon} \text{ and } \frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} - R_1(b) \le \bar{\epsilon}.$$

Choose D_1 and D_2 to make $\frac{\int_{D_2} R_2(i)g(s)f(i)dsdi}{\int_{D_2} g(s)f(i)dsdi}$ as large as possible and $\frac{\int_{D_1} R_2(i)g(s)f(i)dsdi}{\int_{D_1} g(s)f(i)dsdi}$ as close to $R_2(a)$ as possible (indeed the distance is also controlled by $\bar{\epsilon}$), and $\frac{\int_{D_2} R_1(s)g(s)f(i)dsdi}{\int_{D_2} g(s)f(i)dsdi}$ as large as possible. Then it suffices to have

$$\frac{\frac{\int_{D_2} R_2(i)g(s)f(i)dsdi}{\int_{D_2} g(s)f(i)dsdi} - \frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi}}{\frac{\int_{D_0} R_2(i)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} - R_2(a) + \bar{\epsilon}} > \frac{\frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi} - (R_1(\bar{s}) - \bar{\epsilon})}{\frac{\int_{D_1} R_1(s)g(s)f(i)dsdi}{\int_{D_1} g(s)f(i)dsdi} - \frac{\int_{D_0} R_1(s)g(s)f(i)dsdi}{\int_{D_0} g(s)f(i)dsdi}}.$$

Note that both $\frac{\int_{D_2} R_2(i)g(s)f(i)dsdi}{\int_{D_2} g(s)f(i)dsdi}$ and $\frac{\int_{D_2} R_1(s)g(s)f(i)dsdi}{\int_{D_1} g(s)f(i)dsdi}$ can be sufficiently large. In the limit case, we can make

$$\frac{\frac{\int_{D_2} R_2(i)g(s)f(i)dsdi}{\int_{D_2} g(s)f(i)dsdi} - R_2(a)}{2\bar{\epsilon}} > \frac{2\bar{\epsilon}}{\frac{\int_{D_1} R_1(s)g(s)f(i)dsdi}{\int_{D_1} g(s)f(i)dsdi} - R_1(b)},$$

which is always possible. $\square$

Proof of Proposition 7. First, note $f(i|e_o)$ may potentially involve $\bar{I}$, so we do some transformation. Let $\tau = i/\bar{I}$ be the ratio. Suppose $\tau \in [0,1]$ has a distribution $\tilde{F}(\tau|e_o)$, which is independent of $\bar{I}$. Then, $F(i|e_o) = \tilde{F}(i/\bar{I}|e_o)$ and $f(i|e_o) = \frac{1}{\bar{I}}\tilde{f}(i/\bar{I}|e_o)$.

We denote $\tilde{s}(\mu, \lambda, \tau)$ as the cut-off solving $\mu + \sum_j \lambda_j(1 - \frac{g(s|\hat{e}_m^j)\tilde{f}(\tau|\hat{e}_o^j)}{g(s|e_m^*)\tilde{f}(\tau|e_o^*)}) = 1$ for $s < \tau\bar{I}$, where j is the index for the j-th NJ constraint. And denote $\tau_s(\mu, \lambda)$ as the minimum solution for $\tilde{s}(\mu, \lambda, \tau) = \tau\bar{I}$, $\tau_m(\mu, \lambda)$ as the minimum solution for $\mu + \sum_j \lambda_j(1 - \frac{(1 - G(\tau\bar{I}|\hat{e}_m^j))\tilde{f}(\tau|\hat{e}_o^j)}{(1 - G(\tau\bar{I}s|e_m^*))\tilde{f}(\tau|e_o^*)}) = 1$.

By strong duality, and Theorem 2, we have

$$-W(\bar{I}) = \min_{\mu, \lambda} \phi(\mu, \lambda, \bar{I})$$

$$= \min_{\mu, \lambda} \int_{\tau_s(\mu,\lambda)}^1 \int_{\tilde{s}(\mu,\lambda,\tau)}^{\tau\bar{I}} \left(-1 + \mu + \sum_j \lambda_j(1 - \frac{g(s|\hat{e}_m^j)\tilde{f}(\tau|\hat{e}_o^j)}{g(s|e_m^*)\tilde{f}(\tau|e_o^*)}) \right) g(s|e_m^*)\tilde{f}(\tau|e_o^*)dsd\tau$$

$$+ \int_{\tau_m(\mu,\lambda)}^1 \left(-1 + \mu + \sum_j \lambda_j(1 - \frac{(1 - G(\tau\bar{I}|\hat{e}_m^j))\tilde{f}(\tau|\hat{e}_o^j)}{(1 - G(\tau\bar{I}s|e_m^*))\tilde{f}(\tau|e_o^*)}) \right) (1 - G(\tau\bar{I}s|e_m^*))\tilde{f}(\tau|e_o^*)d\tau$$

$$-\sum_j \lambda_j[c(e^*) - c(\hat{e}^j)] - \mu[c(e^*) + \underline{U}].$$

Note the dual is convex in (μ, λ). Let (μ^*, λ^*) be the solution of the dual minimization. Then by the envelope theorem, we have

$$
\begin{aligned}
-\frac{dW(\bar{I})}{d\bar{I}} &= \frac{\partial}{\partial \bar{I}}\phi(\mu^*, \lambda^*, \bar{I}) \\
&= \int_{\tau_s(\mu^*,\lambda^*)}^{1} \tau\left(-1 + \mu^* + \sum_j \lambda_j^*(1 - \frac{g(\tau\bar{I}|\hat{e}_m^j)\tilde{f}(\tau|\hat{e}_o^j)}{g(\tau\bar{I}|e_m^*)\tilde{f}(\tau|e_o^*)})\right)g(\tau\bar{I}|e_m^*)\tilde{f}(\tau|e_o^*)d\tau \\
&\quad - \int_{\tau_m(\mu^*,\lambda^*)}^{1} \tau\left((-1+\mu^*)g(\tau\bar{I}|e_m^*)\tilde{f}(\tau|e_o^*) + \sum_j \lambda_j^*[g(\tau\bar{I}|e_m^*)\tilde{f}(\tau|e_o^*) - g(\tau\bar{I}|\hat{e}_m^j))\tilde{f}(\tau|\hat{e}_o^j)]\right)d\tau \\
&= -\int_{\tau_m(\mu^*,\lambda^*)}^{\tau_s(\mu^*,\lambda^*)} \tau\left(-1+\mu^* + \sum_j \lambda_j^*(1 - \frac{g(\tau\bar{I}|\hat{e}_m^j)\tilde{f}(\tau|\hat{e}_o^j)}{g(\tau\bar{I}|e_m^*)\tilde{f}(\tau|e_o^*)})\right)g(\tau\bar{I}|e_m^*)\tilde{f}(\tau|e_o^*)d\tau \\
&> -\left(-1+\mu^* + \sum_j \lambda_j^*(1 - \frac{g(\tau_s(\mu^*,\lambda^*)\bar{I}|\hat{e}_m^j)\tilde{f}(\tau_s(\mu^*,\lambda^*)|\hat{e}_o^j)}{g(\tau_s(\mu^*,\lambda^*)\bar{I}|e_m^*)\tilde{f}(\tau_s(\mu^*,\lambda^*)|e_o^*)})\right)\int_{\tau_m(\mu^*,\lambda^*)}^{\tau_s(\mu^*,\lambda^*)} \tau g(\tau\bar{I}|e_m^*)\tilde{f}(\tau|e_o^*)d\tau = 0,
\end{aligned}
$$

where the inequality is by MLRP and $\tau_s(\mu^*, \lambda^*) \geq \tau_m(\mu^*, \lambda^*)$.

References

Aliprantis, C.D., K.C. Border. 2006. *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Second Edition. New York, NY: Springer.

Anupindi, R., M. Dada, S. Gupta. 1998. Estimation of consumer demand with stock-out based substitution: An application to vending machine products. *Marketing Sci.* **17**(4) 406–423.

Arya, A., J. Glover, B. Mittendorf. 2007. Aligning incentives by capping bonuses. Antle, R., F. Gjesdal, P. J. Liang, eds. *Essays in Accounting Theory in Honour of Joel S. Demski* (New York, NY: Springer) 169–182.

Atalı, A., H. L. Lee, and Ö. Özer. 2009. If the inventory manager knew: Value of visibility and RFID under imperfect inventory information. *Stanford University Working Paper.*

Baiman, S., S. Netessine, R. Saouma. 2010. Informativeness, incentive compensation, and the choice of inventory buffer. *Accounting Rev.* **85**(6) 1839–1860.

Barvinok, A. 2002. *A Course in Convexity*. Graduate Studies in Mathematics **54**. Providence, RI: American Mathematical Society.

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

Beck, A., C. Peacock. 2009. *New Loss Prevention: Redefining Shrinkage Management.* Basingstoke, United Kingdom: Palgrave Macmillan.

Besbes, O., A. Muharremoglu. 2013. On implications of demand censoring in the newsvendor problem. *Management Sci.* **59**(6) 1407–1424.

Bond, P., A. Gomes. 2009. Multitask principal-agent problems: Optimal contracts, fragility, and effort mis-allocation. *J. Econom. Theory* **144**(1) 175-211.

Chao, G. H., S. M. R. Iravani, R. C. Savaskan. 2009. Quality improvement incentives and product recall cost sharing contracts. *Management Sci.* **55**(7) 1122–1138.

Chen, K.-Y., S. Li, Y. Rong. 2018. The behavioral promise and pitfalls in compensating store managers. *City University of New York Working Paper.*

Chu, L. Y., G. Lai. 2013. Salesforce contracting under demand censorship. *Manufacturing Service Oper. Management* **15**(2) 320–334.

Conlon, C. T., J. H. Mortimer. 2013. Demand estimation under incomplete product availability. *Amer. Econom. J.: Microeconomics* **5**(4) 1–30.

Crittenden, V. L. 1992. Close the marketing/manufacturing gap. *MIT Sloan Management Rev.* **33**(3) 41–52.

Dai, T., K. Jerath. 2013. Salesforce compensation with inventory considerations. *Management Sci.* **59**(11) 2490–2501.

Dai, T., K. Jerath. 2016. Impact of inventory on quota-bonus contracts with rent sharing. *Oper. Res.* **64**(1) 94–98.

Datar, S., S. C. Kulp, R. A. Lambert. 2001. Balancing performance measures. *J. Accounting Res.* **39**(1) 75-92.

DeHoratius, N., A. Raman. 2007. Store manager incentive design and retail performance: An exploratory investigation. *Manufacturing Service Oper. Management* **9**(4) 518–534.

Dewatripont, M., I. Jewitt, J. Tirole. 1999. The economics of career concerns, Part II: Application to missions and accountability of government agencies. *Rev. Econom. Stud.* **66**(1) 199–217.

Feiler, D. C., J. D. Tong, R. Larrick. 2013. Biased judgment in censored environments. *Management Sci.* **59**(3) 573–591.

Feltham, G. A., J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *Accounting Rev.* **69**(3) 429–453.

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

Herweg, F., D. Muller, P. Weinschenk. 2010. Binary payment schemes: moral hazard and loss aversion. *Amer. Econom. Rev.* **100**(5) 2451–2477.

Ho, T. H., C. S. Tang. 2004. Introduction to the special issue on marketing and operations management interfaces and coordination. *Management Sci.* **50**(4) 429–430.

Hollinger, R. C., L. Langton. 2004. National retail security survey: Final report. Report, University of Flordia, Gainesville, FL.

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

Holmstrom, B., P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *J. Law Econom. Organ.* **7**(Sp) 24–52.

Huh, W. T., R. Levi, P. Rusmevichientong, J. B. Orlin. 2011. Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. *Oper. Res.* **59**(4) 929–941.

Jain, A., N. Rudi, T. Wang. 2014. Demand estimation and ordering under censoring: Stock-out timing is (almost) all you need. *Oper. Res.* **63**(1) 134–150.

Innes, R. D. 1990. Limited liability and incentive contracting with ex-ante action choices. *J. Econom. Theory* **52** 45–67.

Jain, S. 2012. Self-control and incentives: An analysis of multiperiod quota plans. *Marketing Sci.* **31**(5) 855–869.

Jerath, K., S. Netessine, Z. J. Zhang. 2007. Can we all get along? incentive contracts to bridge the marketing and operations divide. *University of Pennsylvania Working Paper.*

Jerath, K., S.-H. Kim, R. Swinney. 2017. Product quality in a distribution channel with inventory risk. *Marketing Sci.* **36**(5) 747–761.

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

Ke, R., C. T. Ryan. 2018. Monotonicity of optimal contracts without the first-order approach. *Oper. Res.* **66**(4) 1101–1118.

Kirkegaard, R. 2017a. Moral hazard and the spanning condition without the first-order approach. *Games Econom. Behavior* **102**(3) 373–387.

Kirkegaard, R. 2017b. A unifying approach to incentive compatibility in moral hazard problems. *Theoretical Econom.* **12**(1) 25–51.

Krishnan, H., R. A. Winter. 2012. The economic foundations of supply chain contracting. *Foundations and Trends in Technology, Information and Operations Management* **5**(3–4) 147–309.

Krishnan, J., M. Fisher. 2005. Store level execution at Wawa. Case study, University of Pennsylvania.

Laffont, J. J., D. Martimort. 2009. *The Theory of Incentives: The Principal-Agent Model.* Princeton, NJ: Princeton University Press.

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

Long, X, J. Nasiry. 2018. Wage transparency and social comparison in salesforce compensation. *University of Wisconsin Working Paper.*

Mirrlees, J. A. 1999. The theory of moral hazard and unobservable behaviour: Part I. *Rev. Econom. Stud.* **66**(1) 3–21.

Misra, S., E. J. Pinker, R. A. Shumsky. 2004. Salesforce design with experience-based learning. *IIE Transactions* **36**(10) 941–952.

Nikoofal, M. E., M. Gümüş. 2018. Quality at the source or at the end? Managing supplier quality under information asymmetry. *Manufacturing Service Oper. Management* **20**(3) 498–516.

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

Page, F. H. 1987. The existence of optimal contracts in the principal-agent model. *J. Math. Econom.* **16**(2) 157–167.

Perakis, G., G. Roels. 2010. Robust controls for network revenue management. *Manufacturing Service Oper. Management* **12**(1) 56–76.

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

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

Rogerson, W. P. 1985. The first-order approach to principal-agent problems. *Econometrica* **53**(6) 1357–1367.

Rudi, N., D. Drake. 2014. Observation bias: The impact of demand censoring on newsvendor level and adjustment behavior. *Management Sci.* **60**(5) 1334–1345.

Sethi, S. P., G. L. Thompson. 2000. *Optimal Control Theory: Applications to Management Science and Economics.* , New York, NY: Springer.

Shapiro, B. P. 1977. Can marketing and manufacturing coexist? *Harvard Bus. Rev.* **55**(5) 104–114.

Ton, Z., A. Raman. 2004. The effect of product variety and inventory levels on misplaced products at retail stores: A longitudinal study. *Harvard Business School Working Paper.*

Online Appendix to
"Incentive Design for Operations-Marketing Multitasking"

In this online appendix, we provide technical details and robustness checks complementing the paper. In Section OA.1, we present a numerical example to illustrate the procedure for deriving the optimal mast-and-sail compensation plan. In Section OA.2, we discuss some results on corner compensation plans. In Section OA.3, we present two numerical examples for deriving the optimal corner compensation plan and compare its performance against the optimal mast-and-sail compensation plan. In Section OA.4, we examine the efficiency loss due to demand censoring. In Section OA.5, we derive the analytical results for the case of single-tasking (i.e., the agent manages inventory only) and then compare the inventory thresholds under single-tasking and multitasking. In Section OA.6, we numerically illustrate the effect of the compensation ceiling on the optimal compensation plan and the firm's expected cost of compensating the store manager. In Section OA.7, we discuss an alternative approach to microfound the compensation ceiling.

OA.1:   Numerical Illustration of the Mast-and-Sail Compensation Plan

We now use a numerical example to illustrate an optimal bonus region of the "mast and sail" structure (as seen in Figure 5(a)) in the case of outcome distributions that satisfy the MLRP. This example shows the delicacy of numerical computation in this setting, which is typical of moral hazard problems. Our extensive numerical simulations use similar logic to that found in this example.

EXAMPLE OA.1. Consider the following instance of the multitasking store manager. The distribution functions of operating and marketing effort are $F(i|e_o) = i^{e_o}$ and $G(s|e_m) = s^{e_m}$, respectively, where $e_o \in \{e_o^L, e_o^H\}$ and $e_m \in \{e_m^L, e_m^H\}$ where $e_o^L = e_m^L = 1$ and $e_o^H = e_m^H = 2$. The target action is $(e_o^H, e_m^H) = (2,2)$. The cost function is $c(e_o^H, e_m^H) = 5$, $c(e_o^H, e_m^L) = c(e_o^L, e_m^H) = 4$ and $c(e_o^L, e_m^L) = 2$. The resource constraint for the firm has $\bar{w} = 10$.

For now, we suppose an optimal choice of ω (as guaranteed to exist by Theorem 2) has $\omega_{e_o^L, e_m^L} = 1$ and $\omega_{e_o^H, e_m^L} = \omega_{e_o^L, e_m^H} = 0$. We construct the associated trigger value t below, and also show the resulting compensation plan with these choices of parameters is indeed feasible to (4) and thus optimal.

The condition $R_{e_o, e_m}^{\text{NSO}}(i,s) \geq t$ can be expressed as

$$1 - t \geq \frac{f(i|e_o)g(s|e_m)}{f(i|e_o^H)g(s|e_m^H)} = \frac{e_o}{e_o^H}\frac{e_m}{e_m^H}i^{e_o - e_o^H}s^{e_m - e_m^H}$$
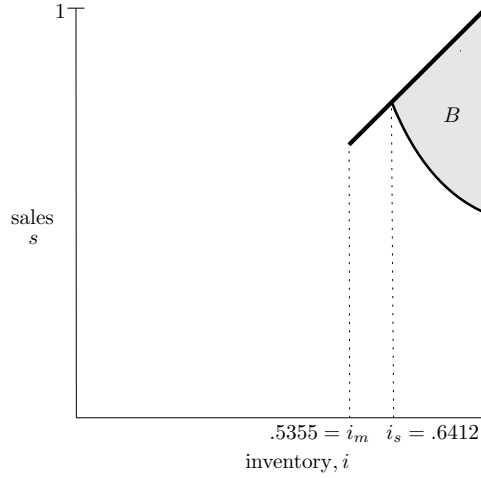
Figure OA.1 **The bonus region of the optimal compensation plan for Example OA.1.**

and the condition $R^{\text{SO}}_{e_o, e_m}(i) \geq t$ amounts to

$$1 - t \geq \frac{e_o i^{e_o - 1}(1 - i^{e_m})}{e^H_o i^{e^H_o - 1}(1 - i^{e^H_m})}.$$

These two conditions are equivalent to (suppose $e_o < e^H_o$)

$$i \geq \begin{cases} \frac{1}{4s(1-t)} & \text{if } s < i \\ i_m(t) & \text{if } s = i \end{cases},$$

where $i_m(t) \triangleq \frac{-1 + \sqrt{1 + 2/(1-t)}}{2}$. Therefore, the bonus region B^{NSO} is

$$\left\{ (i, s) : \frac{1}{4i(1-t)} \leq s < i \text{ and } i \geq i_s(t) \right\},$$

where $i_s(t) \triangleq \frac{1}{2\sqrt{1-t}}$.

The next step is to determine t. Because $\omega_{e^L_o, e^L_m} = 1$, the NJ constraint between (e^H_o, e^H_m) and (e^L_o, e^L_m) is tight. We can isolate for t in the resulting equality to determine t. The tight NJ constraint is

$$\int_{R(i,s|e) \geq t} R(i,s) f(i|e^*_o) g(s|e^*_m) ds di = (c(e^H_o, e^H_m) - c(e^L_o, e^L_m)) \cdot \bar{w}^{-1},$$

which can be rewritten as

$$\int_{i_s(t)}^1 \int_{\frac{1}{4i(1-t)}}^i (4si - 1) ds di + \int_{i_m(t)}^1 (2i(1 - i^2) - (1 - i)) di = 0.3.$$

Solving for t results in $t^* \simeq 0.3919$.

We check whether the trigger compensation plan w^* with $t^* \simeq 0.3919$, $\omega_{e_o^L, e_m^L} = 1$, and $\omega_{e_o^H, e_m^L} = \omega_{e_o^L, e_m^H} = 0$ is an optimal compensation plan. It suffices to check that the remaining no-jump constraints are feasible. We first check that no profitable deviation to $e = (e_o^H, e_m^L)$ exists. Under the trigger compensation plan, the marginal revenue of deviation to (e_o^H, e_m^L) is

$$\bar{w} \int_{i_s(t)}^1 \int_{\frac{1}{4i(1-t)}}^i (4si - 2i) ds di + \bar{w} \int_{i_m(t)}^1 (2i(1 - i^2) - 2i(1-i)) di.$$

Plugging $t^* \simeq 0.3919$ into the above object yields $1.748 > 1 = c(e_o^H, e_m^H) - c(e_o^H, e_m^L)$. Therefore, the store manager will not deviate to (e_o^H, e_m^L).

Similarly, the store manager's marginal revenue of deviation to (e_o^L, e_m^H) is

$$\bar{w} \int_{i_s(t)}^1 \int_{\frac{1}{4i(1-t)}}^i (4si - 2s) ds di + \bar{w} \int_{i_m(t)}^1 (2i(1 - i^2) - (1 - i^2)) di.$$

Plugging $t^* \simeq 0.3919$ into the above object yields $1.864 > 1 = c(e_o^H, e_m^H) - c(e_o^L, e_m^H)$. Therefore, the store manager will not deviate to (e_o^L, e_m^H). The bonus region is illustrated in Figure OA.1. Note that $i_m(t^*) = 0.5355 < i_s(t^*) = 0.6412$, and so we have a mast-and-sail bonus region as plotted in Figure OA.1. Under this optimal compensation plan, the probability of payout under the optimal mast-and-sail compensation plan is 51.96%.

OA.2: Some Results on Corner Compensation Plans

To quantify the performance loss of corner compensation plans, we first need to describe the structure of an *optimal* corner compensation plan. Luckily, the analysis under a corner compensation problem greatly simplifies, as evidenced by the following result.

PROPOSITION OA1. *The expected wage payout of the corner compensation plan (a, b) is $\bar{w}(1 - F(a|e_o^*)(1 - G(b|e_m^*))$ where $(1 - F(a|e_o^*)(1 - G(b|e_m^*))$ is the probability of paying out the bonus, where (e_o^*, e_m^*) is the target effort level to be implemented.*

Proof:

$$\int_a^{\bar{I}} \int_b^i f(i|e_o^*) g(s|e_m^*) ds di + \int_a^{\bar{I}} f(i|e_o^*)(1 - G(i|e_m^*)) di$$
$$= \int_a^{\bar{I}} [G(i|e_m^*) - f(i|e_o^*) G(b|e_m^*)] di + \int_a^{\bar{I}} f(i|e_o^*)(1 - G(i|e_m^*)) di$$
$$= [1 - F(a|e_o^*)][1 - G(b|e_m^*)],$$

completing the proof. □

Given this characterization of expected wage payout, problem (2) evaluated at the corner compensation plan (a, b) becomes (after some basic simplifications):

$$\max_{a,b} \quad r \mathbb{E}[S|e_o^H, e_m^H] - \bar{w}(1 - F(a|e_o^H)(1 - G(b|e_m^H)) \tag{OA1a}$$

$$\text{s.t.} \quad [1-F(a|e_o^H)][1-G(b|e_m^H)]-[1-F(a|e_o^L)][1-G(b|e_m^L)] \geq \frac{c(e_o^H,e_m^H)-c(e_o^L,e_m^L)}{\bar{w}} \tag{OA1b}$$

$$[F(a|e_o^L)-F(a|e_o^H)][1-G(b|e_m^H)] \geq \frac{c(e_o^H,e_m^H)-c(e_o^H,e_m^L)}{\bar{w}} \tag{OA1c}$$

$$[1-F(a|e_o^H)][G(b|e_m^L)-G(b|e_m^H)] \geq \frac{c(e_o^H,e_m^H)-c(e_o^L,e_m^H)}{\bar{w}} \tag{OA1d}$$

$$b \leq a, \tag{OA1e}$$

again recalling that we are in the binary action, MLRP setting with target effort level $(e_o^*, e_m^*) = (e_o^H, e_m^H)$, and we look at the setting where the store manager earns positive rents (as discussed after Proposition 6) and so ignore the IR constraint.

Optimal solutions to (OA1) are relatively easy to characterize, depending on which of the constraints are slack or tight. The next result draws out the implications of certain constraints being tight on the structure of the optimal solution. The proof derives from standard algebraic manipulation of constraints and is thus omitted.

PROPOSITION OA2. *If constraint (OA1b) is tight at optimality, the optimal corner compensation plan (a, b) has a and b satisfy the equation:*

$$\frac{\mathcal{H}^f(a|e_o^H)}{\mathcal{H}^g(b|e_m^H)} = \frac{\mathcal{H}^f(a|e_o^L)}{\mathcal{H}^g(b|e_m^L)}, \tag{OA2}$$

where $\mathcal{H}^f(a|e_o^*) = \frac{f(a|e_o^*)}{1-F(a|e_o^*)}$ is the hazard rate for density f and $\mathcal{H}^g$ is the hazard rate of density g. On the other hand, if (OA1b) does not bind, then $a = b$, where a is characterized by setting either constraint (OA1c) or (OA1d) to be tight.

OA.3:   Numerical Examples of Corner Compensation Plans

To give a sense of how to compute the optimal corner compensation plan we give two concrete examples. These examples refer to some auxiliary results in the same online appendix that help us analyze corner compensation plans.

EXAMPLE OA.2 (EXAMPLE OA.1 CONTINUED). Returning to the set up in Example OA.1, the optimal corner compensation plan solves problem (OA1) and can be simply stated as:

$$\min_{a,b} \quad (1-b^2)(1-a^2)$$

$$\text{s.t.} \quad s.t. \ (1-b^2)(1-a^2)-(1-b)(1-a) \geq= 0.3$$

$$(1-b^2)(1-a^2)-(1-b^2)(1-a) \geq 0.2$$

$$(1-b^2)(1-a^2)-(1-b)(1-a^2) \geq 0.2$$

$$b \leq a.$$

Note we have formulated the objective to minimize the expected probability of paying out the bonus $\bar{w}$ (following Proposition OA1) which is equivalent to the objective function (OA1a).

Now, the optimality condition for the optimal corner compensation plan

$$\frac{\mathcal{H}^f(a|e_o^*)}{\mathcal{H}^g(b|e_m^*)} = \frac{\mathcal{H}^f(a|e_o)}{\mathcal{H}^g(b|e_m)}$$

from equation (OA2) implies

$$\frac{\frac{2a}{(1-a^2)}}{\frac{2b}{(1-b^2)}} = \frac{\frac{1}{(1-a)}}{\frac{1}{(1-b)}},$$

which yields $a = b$. Combined with the other case of Proposition OA2, this implies $a = b$ in optimality. In this case, we can show the optimal choice of ω has $\omega_{e_o^L, e_m^L} = 1$ and $\omega_{e_o^H, e_m^L} = \omega_{e_o^L, e_m^H} = 0$. The optimal corner compensation plan can be obtained through solving $(1 - b^2)(1 - b^2) - (1 - b)^2 = 0.3$, which yields $b^* = 0.5234$. Under the optimal corner compensation plan, the probability of payout is 0.5271. By comparison, the probability of payout under the optimal mast-and-sail compensation plan is 0.5196. So the performance gap is (0.5271-0.5196)/0.5196=1.44%.

EXAMPLE OA.3. Consider an instance with the same setting as in Example OA.1 except that the cost functions are as follows: $c(e_o^H, e_m^H) = 3.4$, $c(e_o^H, e_m^L) = 1.5$, $c(e_o^L, e_m^H) = 1.8$, and $c(e_o^L, e_m^L) = 1$. In this case, we can show the optimal choice of ω has $\omega_{e_o^H, e_m^L} = 1$ and $\omega_{e_o^L, e_m^L} = \omega_{e_o^L, e_m^H} = 0$. Thus, we can solve for b^* by setting $(1 - b^2)(1 - b^2) - (1 - b^2)(1 - b) = 0.19$, which yields $b^* = 0.4896$, which corresponds to a probability of payout of 0.5781. By comparison, the optimal mast-and-sail compensation plan has a probability of payout of 0.5148. Thus, using a corner compensation plan leads to an efficiency loss of $(0.5781 - 0.4976)/0.5148 = 12.30\%$.

OA.4: Fully Observed Demand

A natural benchmark is to look at the scenario where demand is fully observed and not censored by inventory; that is, both the firm and store manager can observe Q. This situation is much simpler than the case of censored demand. The analysis in Section 4 still applies and it can be shown that

$$R_{e_o, e_m}(i, q) = 1 - \frac{f(i|e_o) g(q|e_m)}{f(i|e_o^*) g(q|e_m^*)}, \tag{OA3}$$

which is precisely what was analyzed before as $R_{e_o, e_m}^{NSO}(i, s)$ in (12) for $s \leq i$. Similar reasoning thus yields the result:

PROPOSITION OA3. *A decreasing and continuous function q^f exists such that an optimal contract w^f to the fully observed demand exists with the form*

$$w^f(i, q) = \begin{cases} \overline{w} & \text{if } (i, q) \in B^f. \\ 0 & \text{otherwise,} \end{cases}$$

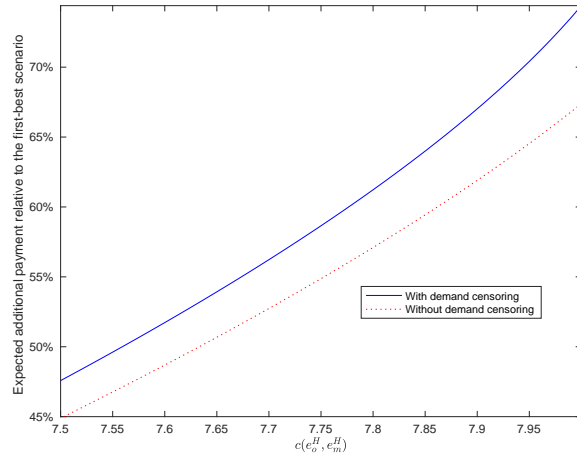*where $B^f = \{(i, q) : q \geq q^f(i)\}$.*

Figure OA.2 **Expected additional payment relative to the first-best scenario with and without demand censoring.**
Parameters: $\bar{w} = 20$, $c(e_o^L, e_m^L) = 1$ **and** $c(e_o^H, e_m^L) = c(e_o^L, e_m^H) = 5$. **We vary** $c(e_o^H, e_m^H)$ **between 7.5 and 8.**

Recall, just as in Proposition 1 for s^*, the domain of the function q^f need not be all of $[0, \bar{I}]$ and an $\bar{i}$ may exist such that $B^f \subseteq [\bar{i}, \bar{I}] \times [0, \bar{Q}]$. Thus, the bonus region of the optimal contract in the fully observed demand case is an "expanded sail" that is not truncated by the 45° line.

Next, we use a numerical experiment to demonstrate the efficiency loss due to demand censoring by comparing the firm's expected additional payment to the store manager—compared to the first-best scenario—for inducing the same effort level. We illustrate our result in Figure OA.2.

OA.5: Single-tasking versus Multitasking

A natural question is to compare a multi-tasking manager to a single-tasking manager responsible for only one of these tasks. In this section, we ask about the optimal incentives for an inventory manager who finds himself in the same situation as the store manager in the main body of the paper, but cannot influence demand through his efforts (i.e., he can only undertake operational effort). Building on our analysis of the general multi-tasking setting in Section 4 (in particular Theorem 2), one can show the optimal contract for the inventory manager problem is an inventory quota-bonus compensation plan.

THEOREM OA.1. *There exist nonnegative multipliers ω and a "target" t such that an optimal solution to the single-tasking inventory manager problem of the following form exists:*

$$w^*(i, s) = \begin{cases} \bar{w} & \text{if } R(i, s) \geq t, \\ 0 & \text{otherwise} \end{cases}, \tag{OA4}$$
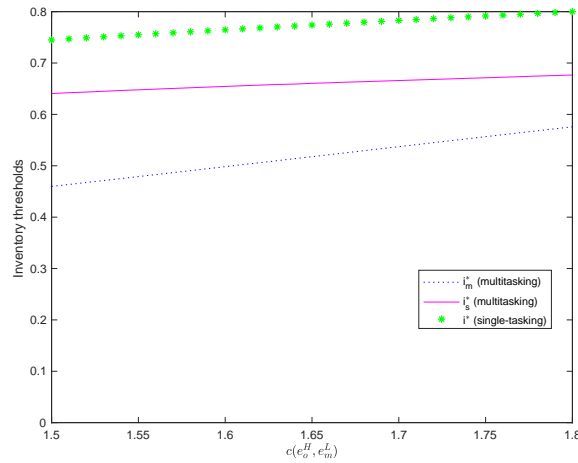
Figure OA.3   **Comparison of inventory thresholds under single-tasking versus multitasking. All the parameters are the same as in Figure 9(b)**

where now

$$R(i,s) = 1 - \frac{f(i|e^L)}{f(i|e^H)},$$

and where the goal is to implement high inventory effort.

Under the MLRP assumption, it is straightforward to show the condition $R(i,s) \geq t$ translates into an inventory threshold i^* where

$$w^*(i,s) = \begin{cases} \bar{w} & \text{if } i \geq i^*. \\ 0 & \text{if } i < i^*. \end{cases} \tag{OA5}$$

Given the optimality of a quota-bonus structure for the inventory management problem, an interesting question is whether the inventory threshold should be higher for the single-tasking agent than the multi-tasking agent. We explore this question numerically. Consider the same set of parameters as in Figure 9(b). In the case of single-tasking, we fix the marketing effort e_m as e_m^H and look for an inventory threshold i^* specified in (OA5). We observe from Figure OA.3 that under single-tasking, the firm consistently chooses an inventory threshold i^* that is higher than both inventory thresholds (i.e., i_s^* and i_m^*) derived from the multitasking setting. In particular, this result implies an inventory manager has a more stringent requirement for earning a bonus in terms of inventory than a multi-tasker who is responsible for inventory (among other things).

OA.6:   Effect of Compensation Ceiling on Compensation Plan and Firm Profitability

We now numerically illustrate the effect of the compensation ceiling ($\bar{w}$) on the design of the optimal compensation plan and the firm's profitability. Consider a case in which all the parameters

are the same as in Figure 9(a) except that (a) we fix $c(e_o^L, e_m^L) = 1.0$ and (b) we vary the value of $\bar{w}$ from 9 to 21. We compute the optimal compensation plans and plot three scenarios in Figure OA.4, corresponding to the cases of $\bar{w} = 9, 15$, and 21, respectively. In addition, we compute the firm's expected cost of compensating the store manager for each combination of parameters. We do not explicitly report the firm's expected profit, which also depends on its unit revenue; a lower compensation cost suggests a higher expected profit. Figure Figure OA.5 shows the relationship between $\bar{w}$ and the firm's expected cost of compensation.
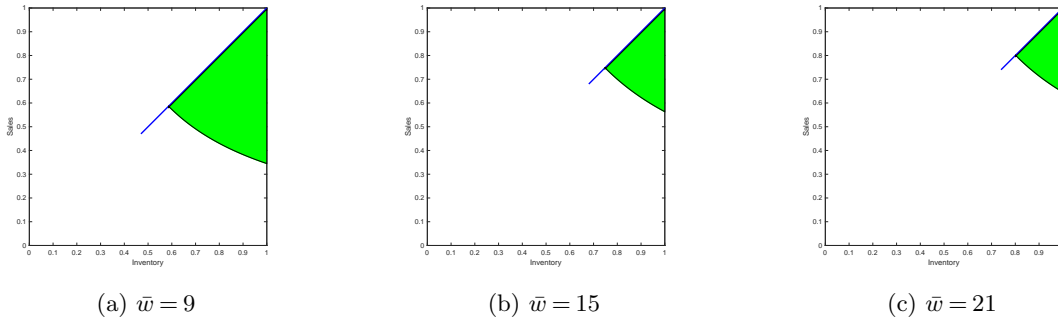

(a) $\bar{w} = 9$ (b) $\bar{w} = 15$ (c) $\bar{w} = 21$

Figure OA.4   **The effect of compensation ceiling on the optimal bonus region. In each panel, the blue line shows the "mast" part of the bonus region, whereas the green area shows the "sail" part.**
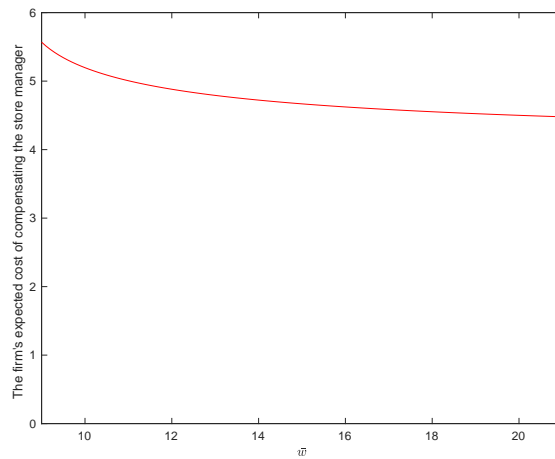

Figure OA.5   **The effect of $\bar{w}$ on the firm's expected cost of compensating the store manager.**
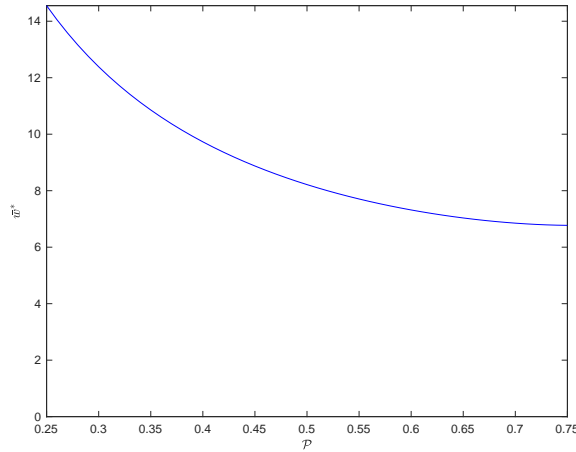
Figure OA.6   The maximum $\bar{w}$ as a function of the agent's minimum likelihood of receiving the bonus ($\mathcal{P}$), which varies between 25% and 75%. All the parameters are the same as in Figure 9(a) except that we fix $c(e_o^L, e_m^L) = 1.0$.

OA.7: An Alternative Approach to Microfound Compensation Ceiling

In Section 8.1, particularly Section 8.1.1, we discuss one approach to microfound the compensation ceiling ($\bar{w}$). In this section, we consider an alternative approach to microfound the compensation ceiling by examining the case in which the agent's probability of receiving the bonus cannot be below a pre-specified threshold (denoted by $\mathcal{P}$). Our analysis is motivated by the observation that in practice, firms sometimes need to ensure each agent, when opting to exert effort, can receive the bonus with a reasonably high likelihood.

Note that for any given $\bar{w}$, we can characterize the optimal compensation package as in Section 5, which gives the optimal solution specified by $t^*(\bar{w})$ such that the bonus region consists of two parts:

$$B^{\mathrm{NSO}}(\bar{w}) \triangleq \left\{ (i,s) \in D^{\mathrm{NSO}} : \sum_{e_o,e_m} \omega_{e_o,e_m} R^{\mathrm{NSO}}_{e_o,e_m}(i,s) \geq t^*(\bar{w}) \right\} \tag{OA6}$$

and

$$B^{\mathrm{SO}}(\bar{w}) \triangleq \left\{ (i,s) \in D^{\mathrm{SO}} : \sum_{e_o,e_m} \omega_{e_o,e_m} R^{\mathrm{SO}}_{e_o,e_m}(i,s) \geq t^*(\bar{w}) \right\}. \tag{OA7}$$

Thus, the problem is equivalent to finding the maximum $\bar{w}$ that satisfies

$$\Pr\left(\sum_{e_o,e_m} \omega_{e_o,e_m} R^{\mathrm{NSO}}_{e_o,e_m}(i,s) \geq t^*(\bar{w})\right) + \Pr\left(\sum_{e_o,e_m} \omega_{e_o,e_m} R^{\mathrm{SO}}_{e_o,e_m}(i,s) \geq t^*(\bar{w})\right) \geq \mathcal{P}. \tag{OA8}$$

Using (OA8) as an additional constraint and endogenously choosing $\bar{w}$, we conduct a numerical experiment and illustrate in Figure OA.6 a sensitivity analysis showing how $\mathcal{P}$ restricts the range of $\bar{w}$ in the incentive-design problem. Our experiment shows that as such a likelihood increases,

as one would expect, the value of $\bar{w}$ decreases. Thus, by incorporating the agent's likelihood of receiving a bonus, this extension can be viewed as a natural way of microfounding $\bar{w}$.