Price Delegation with Learning Agents

Atalay Atasu  
INSEAD, atalay.atusu@insead.edu

Dragos Florin Ciocan  
INSEAD, florin.ciocan@insead.edu

Antoine Désir  
INSEAD, antoine.desir@insead.edu

Many firms delegate pricing decisions to sales agents that directly interact with customers. A premise behind this practice is that sales agents can gather informative signals about the customer’s valuation for the good of interest. The information acquired through this interaction with the customer can then be used to make better pricing decisions. We study the underlying principal-agent problem that arises in such situations. In this setting, the agent can exert costly effort to learn a customer’s valuation and then decide on the price to quote to the customer, while the firm needs to offer a contract to the agent to induce its desired joint learning and pricing behavior. We analyze two versions of this problem: a base model where there is a single customer and a single good, and a generalization where there are multiple customers and limited inventory of the good. For both problems and any additive error $€$, we find contracts that approximate first-best payoffs to within $€$ even if the agent has limited liability, i.e. garners non-negative payments in all states of the world, and shed light on the structure and implementation of such contracts. Under reasonable assumptions, these contracts can be implemented with commissions that are convex-increasing in revenues up to some cap. These contracts continue to perform well under practical adjustments such as commissions with a revenue-sharing structure.

Version: September 22, 2021
Key words: Revenue Management; Pricing; Learning; Sales Force Compensation; Contracting.

Price delegation with learning agents

Atalay Atasu
Technology and Operations Management, INSEAD, atalay.atasu@insead.edu

Dragos Florin Ciocan
Technology and Operations Management, INSEAD, florin.ciocan@insead.edu

Antoine Désir
Technology and Operations Management, INSEAD, antoine.desir@insead.edu

Many firms delegate pricing decisions to sales agents that directly interact with customers. A premise behind this practice is that sales agents can gather informative signals about the customer’s valuation for the good of interest. The information acquired through this interaction with the customer can then be used to make better pricing decisions. We study the underlying principal-agent problem that arises in such situations.

In this setting, the agent can exert costly effort to learn a customer’s valuation and then decide on the price to quote to the customer, while the firm needs to offer a contract to the agent to induce its desired joint learning and pricing behavior. We analyze two versions of this problem: a base model where there is a single customer and a single good, and a generalization where there are multiple customers and limited inventory of the good. For both problems and any additive error $\epsilon$, we find contracts that approximate first-best payoffs to within $\epsilon$ even if the agent has limited liability, i.e. garner non-negative payments in all states of the world, and shed light on the structure and implementation of such contracts. Under reasonable assumptions, these contracts can be implemented with commissions that are convex-increasing in revenues up to some cap. These contracts continue to perform well under practical adjustments such as commissions with a revenue-sharing structure.

version: September 22, 2021

Key words: Revenue management, pricing, learning, sales force compensation, contracting.

1. Introduction

In many commercial settings, firms interact with their customers indirectly, through a pool of salespeople who manage the customer interactions and sales processes. This is almost unanimous in B2B markets, for example in fixed income financial asset or medical equipment sales, but is also common in B2C markets such as real estate and automobile sales. In these settings, salespeople are given price delegation by the firm (Phillips et al. 2015, Lal 1986, Mishra and Prasad 2004, Joseph 2001), meaning that they have decision power over the specific price to quote to a customer. Price delegation recognizes the fact that salespeople can collect idiosyncratic information about customers through their one-on-one interactions, use it to extract valuable signals about those customers’ willingness-to-pay (WTP) and then make better pricing decisions than the firm itself.
A case in point, in the auto-lending industry, Phillips et al. (2015) estimate that delegated prices which are adjusted to customers’ idiosyncratic WTPs improves profit by approximately 11% on average. In such a setting, the firm relinquishes direct control of prices; in turn, the compensation mechanism that it offers to the salesperson becomes a key lever to nudge the salesperson’s decisions towards the firm’s best interests.\footnote{Our interest and the original motivation for this problem comes from an interaction with a revenue management consultancy specializing in pricing solutions in the senior living industry (i.e., retirement homes).}

Theoretical work that explicitly models salespeople who learn through exerting effort and subsequently use this information to price, as we do in the present paper, is scant and we propose to address this gap. Thus, we study a problem where a firm is endowed with a finite inventory of a good.\footnote{Even though inventory considerations are more relevant for product-based sales, we also present, in Section 2, an optimal contract in absence of inventory which can directly be used in a service setting.} There is a finite stream of customers with heterogeneous WTPs drawn independently from a common continuous distribution. The firm delegates pricing to a risk neutral agent (the salesperson), who is capable of learning each customer’s WTP by exerting costly effort; thus, the agent’s informational advantage is endogeneous. When encountering an arriving customer, the agent first decides whether to exert effort and learn, and subsequently decides which price to show to the customer. The firm’s problem is to optimize for contracts maximizing firm payoffs.

We note that agency problems with pricing delegation have long been studied in economics via the principal-agent moral hazard framework (Laffont and Martimort 2009). However, there are a number of important differences between our work and this literature. First, the model of information asymmetry is quite different: the information, which could be thought of colloquially as the “market conditions”, is typically revealed exogeneously and does not depend on the agent’s effort in this stream of literature. In our model, however, information asymmetry is endogeneous, as the agent exerts effort to learn. Second, this literature (for example, Lal 1986, Mishra and Prasad 2004) focuses on demand shifting sales effort where the firm’s profit stochastically increases with the effort level. This relation is typically known and independent of the contract choice. Instead, in our setting, the agent’s pricing decision depends on her learning effort as well as the contract. Consequently, the relationship between the profit distribution and effort is a function of the contract itself. For all these reasons, standard principal-agent approaches fail in our problem.

Our problem is also related to the typical models from revenue management that speak to the optimal pricing of firm inventories, such as the one in seminal paper of Gallego and Van Ryzin (1994, 1997), Maglaras and Meissner (2006) and the subsequent related literature. The typical model there is that the firm has access to population-wide statistics of customer WTP (e.g. a common
distribution that customer valuations are drawn from independently), and centrally optimizes prices with respect to this information. These models do not capture the ability of salespeople to augment this information set with customer level signals and improve on a centrally optimized price.

1.1. Contribution

Our research question is the following: if the firm’s salespeople can be incentivized to acquire information about their customers’ valuations in order to fine-tune their discretionary pricing decisions, when should the firm incentivize this information acquisition and what is the optimal contract to achieve this? Below, we outline our main results addressing this research question.

The single-customer case. We first focus on a “baseline” case, where the firm is endowed with a single good, and there is a single customer with unknown valuation. This focus allows us to analyze a static version of the problem, which leads to a number of fundamental insights about the optimal contract structure to build on for the more general problem.

For this problem, we first verify that a simple contract achieves first best in the absence of limited liability, which is typical in the principal-agent literature. This contract is linear in the (observed) sale price set by the agent and therefore offers a higher payoff when the agent sets a higher price. We next study the implications of a limited liability constraint, which enforces that the contractual payment to the agent is always non-negative. This is an important practical constraint that is encountered in most real world compensation schemes, but complicates the analysis significantly.

In contrast to traditional principal-agent problems where limited liability imposes a strictly positive rent (Laffont and Martimort 2009), we find that there exist contracts that allow the firm to approach the first-best payoff arbitrarily closely. We do so by introducing a novel primal-dual technique that allows us to overcome the technical challenges unique to our problem. In particular, for any $\epsilon > 0$, this technique allows us to identify a contract that generates a payoff which is additively at most $\epsilon$ away from the first-best payoff. We call such a contract an $\epsilon$-optimal contract. Due to limited liability, the structure of this contract is non-linear and intricately depends on the distribution of the customer valuation. If this distribution has increasing hazard rate, the commission structure is non-decreasing and convex in the price up to a certain threshold above which the commission is capped. We believe this is an appealing feature, as it mimics typical salesforce commissions which are convex in sales quantity (Basu et al. 1985, Xiao and Xiao 2020), with the exception that endogenous learning with price delegation requires changing focus from realized sales quantity to realized sales price.
Price-bounded and revenue-share type contracts. Our $\epsilon$–optimal contract has a lottery-type structure where very high but unlikely sales revenues are rewarded with increasingly larger commissions. Accordingly, we also discuss implementations for the single-customer problem that avoid such exceptional rewards and yield contracts further resembling those commonly used in practice. Specifically, we investigate a generalized form of a revenue-share contract, where the fraction shared varies with the sale price, as opposed to a constant revenue-share contract (where the fraction shared is the same regardless of price as in Cachon and Lariviere 2005). This naturally places a bound on how large payments to the agent can grow and can be modeled as an additional constraint in our baseline single-customer problem, which we again tackle via our primal-dual technique. While the firm may not be able to approximate first best to arbitrary precision, in numerical simulations we show that the agent’s rent is significantly lower than the rent achieved by the best constant revenue-share contract. Thus, this contract adjustment is a tempting middle ground between the contracts that are already established in practice, and those that our theoretical analysis directly produces.

The multiple-customer case with inventory. Motivated by the practicalities of the commercial applications mentioned above, we also study a generalized version of the model with limited liability, in which the firm is subject to inventory limitations and there are multiple customers who arrive in sequence. We overcome the challenges inherent in the ensuing dynamic principal-agent model by solving a fluid version of the dynamic problem, where the inventory constraint that the agent’s learning and pricing decisions must satisfy holds in expectation. This allows us to generalize our primal-dual technique to yield an $\epsilon$-optimal contract. This contract is simple in the sense that it decomposes into a series of static single-customer contracts that depend on the initial inventory level and the customer pool size. These differ structurally from the $\epsilon$-optimal single-customer contracts only through the addition of a reservation value on the sale price the agent achieves; we show that carefully choosing this reservation value is sufficient to correctly control the depletion of inventory.

The rest of the paper is organized as follows. Section 1.2 reviews the literature related to this work. Section 2 describes the baseline optimal contracting problem where a single customer may purchase a single good. We give an $\epsilon$-optimal solution to this problem in Section 3. Then, in Section 4, we extend the model, together with our $\epsilon$-optimal contract, to multiple customers and multiple units of inventory. All proofs can be found in the Appendix.
1.2. Literature review

This paper is related to several streams of research detailed below.

Economics. Principal-agent moral hazard models are a cornerstone of economic theory, and we review a few seminal papers here in the context of our research question. Mirrlees (1999) investigates moral hazard for insurance contracts. Given certain assumptions, Mirrlees (1999) finds a family of contracts which approaches the first best. The construction of this result hinges on building heavy agent punishments for certain unlikely outcomes into the contracts. Grossman and Hart (1992) start with a similar model but assume that the agent’s utility function is separable in effort and payment which enables the problem to be cast as a convex program. First best may not always be achievable, but when the agent’s utility function is linear in payment, a linear contract attains it. This is true in our case as well, in the absence of limited liability. Unlike these previous papers, Innes (1990) and Kim (1997) add in limited liability and show that first best can still be achieved under certain conditions. Balmaceda et al. (2016) provide welfare loss bounds from limited liability. Finally, Carroll (2015) shows that linear contracts are optimal for a robust formulation of the principal-agent problem.

There are some key differences between our work and this literature. As noted in the introduction, our learning effort is structurally different from the effort considered there, which is of the demand-shifting type. In the demand-shifting effort case, the workhorse tool used to analyze and to enable first-best achievability is the monotone likelihood ratio property or MLRP (Mirrlees 1999, Grossman and Hart 1992, Innes 1990), or some other statistical property linking effort to the distribution of firm profits; specifically, MLRP informally enforces that higher profits observed by the firm indicate a higher likelihood that effort was exerted. Crucially, for demand-shifting effort, the profit distribution depends on agent effort but not on the contract. In our problem setting, however, this does not appear to be possible. The multi-dimensional and contingent nature of our agent action space, where effort is combined with a pricing decision that depends on effort and the contract function itself, introduces an intermediate layer wherein learning effort does not directly affect the profit distribution. Instead, the learning effort affects the agent’s pricing decision, which in turns affects the profit distribution. Consequently, since the pricing decision of the agent depends on the firm’s contract, the relationship between the profit distribution and effort is a function of the contract itself. This precludes any natural MLRP property for our problem and requires a different technical approach, even in the simpler setting of the single-customer model. We address this challenge by a novel primal-dual approach applied to an auxiliary problem which avoids making strong structural assumptions about how the distribution of firm payoffs depends on effort levels.
Multi-tasking principal-agent models with multiple types of agent effort (Holmstrom and Milgrom 1991) also connect to our work where the agent action space is multi-dimensional (i.e., learning effort plus pricing). However, multi-tasking substantially differs in that it assumes independent outputs for each effort dimension (Bond and Gomes 2009, Dai et al. 2021), whereas in our model pricing depends on effort and, additionally, is observable.

The multiple-customer problem with inventory we analyze relates to dynamic principal-agent models where effort is exerted over time, and the compensation at any such point may depend on past history. Generally, analyzing dynamic principal-agent problems is challenging. One related stream of literature shows that the time dimension of the problem requires more sophisticated, “long-term” contracts which are memoryful, e.g. they use history as input to the current period’s compensation rule; see Radner (1981), Rubinstein (1979), Rogerson (1985). The other stream, which is closer in intent to our problem, shows that simple contracts which describe a single period’s interaction are sufficient, such as the linear contract optimality result of Holmstrom and Milgrom (1987). In a recent paper, Chassang (2013) develop a technique for finding good limited liability dynamic contracts: their approach is to solve for a certain class of simple contracts in the absence of limited liability, and then find limited liability contracts that robustly approximate them.

Our dynamic version of the problem is significantly different from those mentioned above due to the addition of the inventory constraint. This constraint temporarily correlates all the agent-customer interactions, with the agent having to account for the future value of inventory in their current best-response decision. In order to surmount these challenges that arise in the ensuing dynamic mechanism design problem, our technique focuses on a fluid version of the model borrowed from the revenue management literature which we will review below. To the best of our knowledge, we are the first to use a fluid formulation to tractably analyze a dynamic principal-agent problem.

**Pricing and revenue management.** Pricing has been an extremely active area of research in operations, albeit with perhaps less focus on the effects of salesperson incentives. The seminal paper of Gallego and Van Ryzin (1994) considers a firm who prices a finite goods inventory to sequentially arriving customers, whose valuations come from a known distribution. A static price obtained through solving a fluid version of the dynamic problem is shown to be asymptotically optimal in a regime where both the density of customers and inventory grow large. The fluid analysis method of Gallego and Van Ryzin (1994) has often been used in subsequent revenue management work and we defer to Talluri et al. (2004) and Gallego et al. (2019) for a comprehensive literature review. It is a powerful tool to simplify the analysis of dynamic problems, which we employ in our work to obtain results for the multiple-customer problem.
A key feature of Gallego and Van Ryzin (1994) is that the fluid optimal price is obtained by optimizing with respect to the entire distribution of customer WTPs; our model can be viewed as a generalization where the agent has the choice of using this fluid price which is optimized with respect to the entire population’s distribution, versus learning the realized valuation of an individual customer and tailoring the price.

Another related paper to ours is Bhandari and Secomandi (2011). There, a seller repeatedly sells to buyers through a bargaining mechanism, and uses sales data to learn the opportunity cost of their inventory. Our model differs in that we acknowledge this learning may not happen directly in practice, but rather through a salesforce; the corresponding addition of the principal-agent dimension brings forth different challenges to be addressed.

It is also important to note that in the revenue management literature, learning typically carries a different meaning from the one we focus on. Specifically, there is extensive work on statistical learning while pricing, where the firm can use sales information as samples to learn the distribution of customer WTPs it is facing (see Gallego et al. 2019) and refine prices as it learns. In this paper, we do not learn the distribution of customer WTP but assume it is known to the firm and the agents; instead our agents can exert effort to learn individual realizations of customer valuations.

**Operations and marketing interface.** Recognizing the importance of salesforce incentives in commerce, both of these communities have contributed to the literature on moral hazard and price discretion. Examples of optimal contracting work with sales effort are Basu et al. (1985), Lal and Staelin (1986), Rao (1990), Oyer (2000), Chu and Lai (2013). Subsequently, Dai and Jerath (2013, 2016), Xiao and Xiao (2020), Dai et al. (2021) revisit this problem in the presence of potential inventory stock-outs and demand censoring. Lal and Srinivasan (1993), Chen (2000) add dynamic models where the agent’s effort decisions are exerted over time and are state-dependent. On the moral hazard side, the majority of this literature has considered demand shifting sales effort, which lifts the customer’s WTP, but does not endogeneously control the amount of information available to the agent before making the pricing decision. As hinted to above, our learning effort allows the agent to control the customer information they acquire, but on the other hand does not affect WTP per se as sales effort would.

The interaction of sales effort and discretionary pricing has also been an active topic of research; the main tension there is that in the absence of properly set-up incentives, the agent can substitute costly sales effort with price discounting. The first such result is Weinberg (1975), who show that when there is information symmetry and the agent is paid a commission on gross margin, their optimally chosen discretionary prices maximize both the principal’s and agent’s payoffs.
Subsequently, Lal (1986) and Mishra and Prasad (2004) consider the asymmetric case where the agent is endowed with a market signal unavailable to the principal. Lal (1986) shows that if the signal is revealed after contracting and exerting sales effort, the principal can increase it profits by allowing price discretion. On the other hand, if the market signal precedes these, centralized pricing inflicts the principal no loss (Mishra and Prasad 2004). Joseph (2001) further explore the possibility that price discounting can be used by the agent to substitute costly sales effort, and show that placing limits on price discretion can in some cases increase principal profits. We differ from this stream by observing that the agent may need to exert effort to obtain the informational advantage, which requires a fundamentally different optimal contract structure.

Complementing the theory work above, several papers (Stephenson et al. 1979, Frenzen et al. 2010, Homberg et al. 2012) have analyzed sales data in various industries to empirically ascertain the connection between degree of price discretion and firm profitability; their findings have however been mixed and sensitive to the specific setting of the study. A more recent empirical paper (Phillips et al. 2015) considers the relative merits of delegated prices which are adjusted to customers’ idiosyncratic WTPs, versus list prices which are centrally optimized based on directly observable customer factors. Using auto lending data, the authors find that allowing price delegation alone increases profitability, but that even higher gains can be obtained by using delegation in conjunction with centrally optimized list prices. Although these papers acknowledge information acquisition as a key dimension of the selling process, we are not aware of any theoretical or algorithmic work that tackles the optimal contracting problem. In fact, the findings in these papers hold for a given incentive mechanism that may be sub-optimal and under-estimate the true gains from price discretion under learning, making the optimization problem even more important to answer.

2. Single-customer model

2.1. Price discretion and learning effort

Consider a risk-neutral firm selling a product to a customer whose willingness-to-pay $V$ is a continuous random variable with density $f(v)$ and cumulative distribution function $F(v)$, with support $(0, \infty)$. The firm employs a risk-neutral salesperson, henceforth referred to as the agent, who directly interacts with the customer and can, upon exerting some costly binary learning effort $e$, learn the realization of $V$. The agent can either learn the willingness-to-pay $V$ of the customer, i.e. $e = 1$, or decide not to learn anything, i.e. $e = 0$.\(^\text{3}\) Learning is costly to the agent with dis-utility $\psi(e)$,

\(^\text{3}\)In Section 4, we allow the agent to have a randomized strategy where she can decide to learn only with some probability. However, as an optimal deterministic strategy always exists for the agent in the single-customer model, we focus on deterministic strategies to simplify the exposition.
which normalize such that $\psi(0) = 0$ and $\psi(1) = \psi > 0$. The agent’s effort decision is not observable to the firm.

The firm delegates pricing to the agent, where the agent decides on a price $p \in \mathbb{R}^+ \cup \{\infty\}$ which incorporates the information the agent has on the customer’s WTP. We assume that the agent knows the underlying distribution $f(v)$ from which the willingness-to-pay $V$ is drawn. If the agent exerts effort to learn the customer’s willingness-to-pay $V$, she can subsequently use this information to decide on the price $p$. In this case, the agent sets the price $p$ to be a function $\theta(V)$. If the agent does not exert effort to learn, she chooses a constant price $\phi$. We refer to $\theta(V)$ as the informed price policy and to $\phi$ as the uninformed price. In sum, the agent’s action is multi-dimensional and consists of two components: learning (or not) and pricing.

We denote the sale revenue $Y$, which is equal to the price set by the agent in case the sale happens and to 0 in the absence of a sale. More precisely, $Y = p \cdot 1\{V \geq p\}$. As the agent’s learning effort $e$ is not observable by the firm, a contract can only be based on the observable $Y$. For any realization of $Y$ equal to $y \geq 0$, let $t(y)$ be the transfer from the firm to the agent. We also refer to the function $t(\cdot)$ as the contract. Additionally, we assume that $t(y)$ is non-decreasing in $y$ for all $y \geq 0$ and that $E[t(V)] < \infty$ without loss of generality. We denote by $\mathcal{T}$ the space of allowable contracts. We further follow the standard convention (e.g. Laffont and Martimort 2009) that the agent’s utility is separable between money and effort, $U = t(Y) - \psi(e)$. The firm’s payoff is $J = Y - t(Y)$.

### 2.2. Incentive feasible contracts

We focus on the case where the firm wants to induce the agent to exert effort to learn and needs to decide which incentive contract should be used to achieve this outcome. In particular, since the firm looks for a contract that induces effort $e = 1$, the incentive constraint of the agent can be written as

$$\max_{\theta(\cdot)} \left\{ \mathbb{E}[t(\theta(V)) \cdot 1\{V \geq \theta(V)\} + t(0) \cdot 1\{V < \theta(V)\} - \psi] \right\} \geq \max_{\phi > 0} \left\{ t(\phi) \cdot \bar{F}(\phi) + t(0) \cdot F(\phi) \right\},$$

where $\theta(\cdot)$ and $\phi$ are the agent’s pricing decisions as defined above, the left-hand side is the agent’s utility under $e = 1$ and the right-hand side is the same under $e = 0$. For any $t \in \mathcal{T}$, if the agent learns, it is always optimal to price exactly at the value on the almost sure event that the valuation is positive, i.e.

$$\theta(V) = V. \quad (2)$$

---

4 We assume $p > 0$ to avoid having to distinguish between no-sale, i.e. $V < p$, and a sale at 0, i.e. $p = 0$. Additionally, we allow a price $p = \infty$; irrespective of valuation, showing this price to a customer leads to a no-purchase.

5 While the characterization of $\theta(V)$ is relatively straightforward in the single-customer case as we explain below, we keep this general notation for ease of connection to the multiple-customer case in Section 4.
Indeed, if the agent prices strictly above \( v \), then the agent does not sell and receives \( t(0) \leq t(v) \). On the other hand, if the agent uses some price \( v' < v \), then she leaves \( t(v) - t(v') \) on the table which is a non-negative quantity since \( t(\cdot) \) is non-decreasing.\(^6\) We consequently let the informed price policy of the agent satisfy (2) which then implies that the incentive constraint (1) can be simplified as

\[
\mathbb{E}[t(V) - \psi] \geq \max_{\phi > 0} \{t(\phi) \cdot F(\phi) + t(0) \cdot F(\phi)\}. \tag{1C}
\]

In the above constraint (1C), the left-hand side is the expected utility the agent gets when exerting effort to learn and pricing according to the informed price policy set to \( \theta(V) = V \). The right-hand side corresponds to the maximum expected utility the agent can get without exerting learning effort and setting the best uninformed price \( \phi \). Normalizing the agent’s reservation utility at zero, the agent’s participation constraint can be written as

\[
\mathbb{E}[t(V) - \psi] \geq 0. \tag{participation}
\]

We finally impose the following limited liability constraint which ensures that the transfers to the agent are always non-negative (Oyer 2000, Dai and Jerath 2013):

\[
t(p) \geq 0, \quad \forall p. \tag{limited liability}
\]

Putting everything together, the single-customer firm’s problem (SC) can consequently be written as

\[
\begin{align*}
\max_{t \in \mathcal{T}} & \quad \mathbb{E}[V - t(V)] \\
\text{s.t.} & \quad \mathbb{E}[t(V) - \psi] \geq \max_{\phi > 0} \{t(\phi) \cdot F(\phi) + t(0) \cdot F(\phi)\} \\
& \quad \mathbb{E}[t(V) - \psi] \geq 0 \\
& \quad t(p) \geq 0, \quad \forall p.
\end{align*} \tag{SC}
\]

The action space of the agent includes both the learning decision as well as the pricing decision. In particular, when not learning, the agent can choose any price \( \phi \) which makes the action space continuous in nature. This structure for the agent’s action space implies the non-traditional incentive constraint (1C) in the above formulation – this is due to the interplay between the effort decision and the pricing policy which is contingent on it. Moreover, note that our incentive problem focuses on inducing agent learning. This is also fundamentally different from inducing demand

\(^6\) Note that \( \theta(V) = V \) may not be the unique optimal agent strategy unless the contract is strictly increasing. To simplify the exposition, we assume that the agent breaks ties in favor of the informed price policy given in (2). It is however possible to guarantee uniqueness and avoid this issue altogether by refining \( \mathcal{T} \) to only include strictly increasing contracts. In this paper, we do not impose that the contracts are strictly increasing. However, we can always slightly perturb the contracts we study to ensure that (2) is the unique optimal agent strategy.
shifting effort (e.g., Basu et al. 1985, Oyer 2000, Chu and Lai 2013). The typical approach used in that stream of literature is a first-order approach which often requires assumptions such as a stochastic ordering of actions or more particularly MLRP property. Due to the two-dimensional and non-continuous nature of our problem, such assumptions are not applicable in our setting. We discuss this in more details in Section 3.3 and Appendix EC.3. This implies that solving the combined pricing and learning problem with agency needs a rather different treatment which, to the best of our knowledge, has not been addressed to date. Before analyzing the single-customer problem defined in (SC), we discuss a first-best benchmark to set the stage.

2.3. First best implementation in the absence of limited liability

Consistent with the moral hazard literature, we first consider a benchmark where the firm has complete information. In this case, the firm can observe the agent’s effort so that there is no moral hazard problem. We let $J_{FB}$ denote the firm’s first-best payoff. If the firm wants to induce effort (i.e. $e = 1$), its problem can be written as

$$J_{FB} = \max_{t \in T} \{E[V - t(V)] : t(\cdot) \text{ satisfies (participation)}\}.$$  

Any contract that satisfies $E[t(V)] = \psi$ solves the above optimization problem and the firm’s first-best payoff is given by $J_{FB} = E[V] - \psi$. Since we are interested in settings where agency for learning is profitable for the firm, we make the following assumption for the rest of the paper.

**Assumption 1.** $J_{FB} \geq \max_{\phi} \{\phi \cdot \tilde{F}(\phi)\}$. 

In other words, we assume that $J_{FB}$ is always higher than the payoff under the best uninformed price. Note that otherwise, if inducing agent learning is not a dominant strategy, then the firm can always offer a forcing contract by letting $t(p) = \epsilon \cdot 1\{p = \phi^*\}$ for some small $\epsilon > 0$ where $\phi^* = \arg \max_{\phi} \phi \cdot \tilde{F}(\phi)$.

When both the firm and the agent are risk-neutral, it is well-known that delegation in the absence of limited liability is costless to the firm (Mas-Colell et al. 1995, Laffont and Martimort 2009). We show that this is also the case for our problem. In particular, consider the simple linear contract defined for all $p \geq 0$ as

$$t_{lin}(p) = p - J_{FB}.$$  

We next show that this simple linear contract\(^7\) achieves the first-best payoff $J_{FB}$.

\(^7\) Note that there may exist other optimal contracts.
Lemma 1. *The linear contract* $t_{\text{lin}}(\cdot)$ *satisfies (participation) and (IC). Moreover, $\mathbb{E}[V - t_{\text{lin}}(V)] = J_{\text{FB}}^*$, i.e. this contract achieves the first-best payoff.*

The linear contract $t_{\text{lin}}(\cdot)$ is increasing with $p$, so the agent’s commission increases with the value of the sales. Note that $t_{\text{lin}}(\cdot)$ takes a negative value for all $p \leq J_{\text{FB}}^*$. In these cases, the agent must pay the firm which is leveraging negative payments to enforce incentive compatibility. Another way to interpret this linear contract is that the agent makes an upfront payment of $J_{\text{FB}}^*$ to the firm and then keeps all the revenue generated by the sale. In this reasoning, the agent’s risk-neutrality is key, and inefficiencies (in the form of rents) typically appear when this no longer holds. One way to model this is to let the agent be risk-neutral but only for positive outcomes and to impose a limited liability constraint, i.e. impose that the transfers are non-negative. This is what we explore next.

3. *$\epsilon$--optimal contracts in the presence of limited liability*

In this section, we analyze the single-period agency problem defined in (SC). In particular, building on Lemma 1, we study how the presence of the limited liability constraint impacts the firm’s payoff. We first show (Section 3.1) that there is no contract that can attain the first-best payoff while satisfying (IC) and (limited liability) simultaneously. On the positive side, we show that the first-best payoff can be approached arbitrarily closely. In particular, we present a family of bounded contracts in Section 3.2 that contains an $\epsilon$-optimal contract for any $\epsilon > 0$ (Section 3.3) where for any $\epsilon > 0$, we say that a contract $t \in \mathcal{T}$ is $\epsilon$-optimal if $\mathbb{E}[V - t(V)] \geq J_{\text{FB}}^* - \epsilon$.

3.1. *Non implementability of first best in the presence of limited liability*

In the presence of limited liability, we first show that no contract can achieve $J_{\text{FB}}^*$ while simultaneously satisfying (IC) and (limited liability).

**Corollary 1.** *Let $t \in \mathcal{T}$ be such that $\mathbb{E}[t(V)] = \psi$, then (IC) and (limited liability) cannot be both satisfied.*

We also note here that Corollary 1 is stronger than formalized above. In fact, this statement would stand even if we were to relax the requirement that $t(\cdot)$ is non-decreasing. Since the first-best payoff cannot be attained, this implies that limited liability imposes a strictly positive rent as it is often the case in moral hazard problems with limited liability (Mas-Colell et al. 1995, Laffont and Martimort 2009). Surprisingly though, we show in our case that the first-best payoff can be approached arbitrarily closely. Intuitively, instead of imposing negative payments towards the agent, the firm can tie increasingly higher payments to progressively higher price sales.
As a first-step to analyze (SC) and prove the existence of $\epsilon$-optimal contract, we first show that under (limited liability), there always exists an optimal contract that does not compensate the agent in the absence of a sale, i.e. such that $t(0) = 0$. This property, which we formalize in the next lemma, allows us to simplify (IC).

**Lemma 2.** Consider a feasible contract $t(\cdot)$, i.e. which satisfies (participation), (IC) and (limited liability). If $t(0) > 0$, there exists another feasible contract $\tilde{t}(\cdot)$ such that $\tilde{t}(0) = 0$ and $E[V - \tilde{t}(V)] \geq E[V - t(V)]$.

In the rest of the section, we therefore assume that $t(0) = 0$ and focus on finding the right contract $t(p)$ for all other possible sale prices $p > 0$. Note that (IC) now becomes

$$E[t(V) - \psi] \geq \max_{\phi > 0} \{t(\phi) \cdot \bar{F}(\phi)\}.$$ 

We next give some intuition for the underlying structure of the optimal contract. Informally, this structure is to make the right-hand side of the above constraint constant, i.e. make the agent indifferent between all possible uninformed prices $\phi$. Indeed, if the quantity $t(\phi) \cdot \bar{F}(\phi)$ is not constant for all $\phi$, this implies that we can slightly reduce $t(\cdot)$ around all $\phi_0 \in \text{arg max} \{t(\phi) \cdot \bar{F}(\phi)\}$ while maintaining incentive compatibility and at the same time reducing the firm’s expected payment by a small amount. Consequently, an optimal contract should be such that $t(\phi) \cdot \bar{F}(\phi)$ is constant for every possible $\phi$. This implies that $t(\phi)$ should be proportional to $1/\bar{F}(\phi)$. Under such a contract, we have $1/\bar{F}(\phi) \to \infty$ when $\phi \to \infty$, capturing the fact that the firm needs to use arbitrary large rewards to achieve first best.

In the following section, we will formalize the above intuition. A technical challenge we overcome is that making $t(\phi)$ proportional to $1/\bar{F}(\phi)$ for all $\phi$ leads to an ill-defined contract: indeed, it is easy to see that for any continuous distribution $F(\cdot)$, the expected contract payment is proportional to $E[1/\bar{F}(\phi)] = \infty$. Thus, to formalize the above intuition, we need to limit the magnitude of rewards the firm could pay for high-sale prices, while still retaining the proportionality for as much of the domain of $\phi$ as possible. We accomplish this by defining a sequence of auxiliary problems which resemble (SC) except for an additional constraint which bounds the magnitude of $t(\cdot)$. We define a family of contracts which are optimal for these auxiliary problems and then understand the performance of these contracts with respect to (SC) itself.

### 3.2. A family of bounded contracts

To formally analyze the problem, we introduce a family of auxiliary problems where we impose an additional upper bound on the contract’s payment to the original problem (SC). More precisely,
for every $M > 0$, we consider the following constraint that limits the payments the firm can make to the agent:

$$t(p) \leq M, \quad \forall p.$$  \hfill (M-bounded)

Building on the previous intuition of making the agent indifferent between all possible uninformed price $\phi$, we define the following contract for a given $M$ and a threshold $\mu$:

$$t^M_\mu(p) = \min \left\{ M \cdot \frac{\bar{F}(\mu)}{\bar{F}(p)}, M \right\}, \quad \forall p. \quad (3)$$

Even though we use this bound for analysis purposes, note that imposing an upper bound on the agent compensation is common in the contract theory literature and there are natural motivations for considering such a constraint (Bond and Gomes 2009, Dai et al. 2021).

The contract is such that $t^M_\mu(p) \cdot \bar{F}(p)$ is constant for all $p \leq \mu$. For $p \geq \mu$, the contract is constant and equal to the upper bound $M$. Both $M$ and $\mu$ are parameters which we subsequently select so as to obtain our results; as we shall see, in the sequel we shall set $\mu$ as a function of $M$, and then let $M$ tend to infinity to obtain our $\epsilon$-optimality result.

Since $\bar{F}(p)$ is decreasing in $p$, another way to write the contract is as follows:

$$t^M_\mu(p) = \begin{cases} 
M \cdot \frac{\bar{F}(\mu)}{\bar{F}(p)}, & \text{if } p \leq \mu, \\
M, & \text{if } p > \mu.
\end{cases}$$

We next highlight a useful structural property of these contracts.

**Corollary 2.** If the valuation distribution satisfies the monotone hazard rate property, i.e. $f(p)/\bar{F}(p)$ is non-decreasing in $p$, then for all $\mu$ and $M$, $t^M_\mu(p)$ is convex increasing in $p$ for $p \leq \mu$.

When the valuation distribution satisfies the monotone hazard rate property, the contract has an intuitive convex structure whereby the agent’s commission increases convexly with the price. We now show that there always exists an optimal contract of the form $t^M_\mu(\cdot)$ when constraint (M-bounded) is added.

**Proposition 1.** There exists $\tilde{M}$ such that for every $M \geq \tilde{M}$, there exists a threshold $\mu(M)$ such that $t^M(\cdot) = t^M_{\mu(M)}(\cdot)$ solves

$$\max_{t \in T} \{ \mathbb{E}[V - t(V)] : t(\cdot) \text{ satisfies (participation), (IC), (limited liability), (M-bounded)} \}. \quad (4)$$
A common technique in principal-agent model (Laffont and Martimort 2009) is to compute the minimum expected cost of inducing the agent to take a particular action. Instead, in our case, we show that the problem of finding an incentive feasible contract that induces the agent to learn at a particular expected cost (i.e., holding the left-hand side of (IC) fixed) can be formulated as a continuous linear program. Additionally, we provide a closed-form expression for this problem using a primal-dual approach. More precisely, we construct the Lagrangian dual of this continuous linear program and exhibit a pair of primal and dual solutions that attain the same objective value. These solutions are therefore optimal for their respective problem by weak duality. Since this result is constructive, the proof in Appendix EC.1.4 also yields an algorithm to compute $t^M(\cdot)$ via linear programming; we comment further on this in Appendix EC.1.4.

3.3. $\epsilon$-optimal contracts

While in the above section we are able to show that our family of contracts is optimal for the auxiliary problem which enforces (M-bounded), our ultimate goal is to relate these contracts to the original problem (SC). We next show that when $M \to \infty$, we have $\mu(M) \to \infty$ and the rent incurred by $t^M(\cdot)$ vanishes. In other words, if one does not restrict the magnitude of the contract’s payment, limited liability comes for free and our family of contracts approaches first best arbitrarily closely, as we show in the following.

First, for all $M$, let $R(M)$ be the rent imposed by the constraints (limited liability) and (M-bounded). Specifically, for all $M$, the rent captures the difference in payoffs to first best and is defined as

$$R(M) = J^{FB} - \mathbb{E}[V - t^M(V)] = \mathbb{E}[t^M(V)] - \psi.$$ 

Note that $R(M) \geq 0$ by (participation). The key result of this section is that:

**Proposition 2.** $\lim_{M \to \infty} R(M) = 0$.

Furthermore, the following is an immediate consequence of Proposition 2 and formalizes the idea that the first-best payoff can be approached arbitrarily closely even in the presence of limited liability.

**Corollary 3.** For every $\epsilon > 0$, there exists $M$ such that $t^M(\cdot)$ is $\epsilon$-optimal and satisfies (participation), (IC) and (limited liability).

Corollary 3 is a striking result as first best is typically not achievable in moral hazard problems. To the best of our knowledge, the most general condition which has been used in the literature to obtain
first-best performance is the MLRP condition (Innes 1990). As mentioned in our introduction, the multi-dimensional and contingent structure of our agent’s action space make such a condition untenable for our model since the firm’s profit distribution depends on both effort and the contract itself through the price set by the agent (see Appendix EC.3 for a more detailed discussion). This hints at a key difference between the learning effort studied here, and sales effort discussed in our literature review, which lifts customer demand but does not affect the agent’s subsequent actions.

Although the agent’s flexibility to price precludes us from leveraging MLRP-like conditions, we exploit this flexibility to approach first best: in order to show that $R(M)$ vanishes, we leverage the fact that the firm can tie increasingly higher payments to progressively higher price sales. As the probability of these events goes to zero, due to the continuous nature of the distribution, the firm can approach first best. It is also important to note that our results continue to hold even if the valuation distribution has bounded support as long as the distribution is continuous. For example, Figure 1 illustrates the convergence of the rent $R(M)$ as a function of $M$ for two examples of valuation distribution, one unbounded and the other bounded. Moreover, we crucially need the upper bound $M$ to go to $\infty$ in order to approach $J^{FB}$. In case of a discrete distribution, however, limited liability may incur a strictly positive rent as we discuss in the next section.

Finally, the machinery that we develop in this section is not only useful to show the existence of $\epsilon$-optimal contracts. Indeed, we build on the shape of the bounded contracts $t^M_{\mu}(\cdot)$ and importantly on the primal-dual proof of Proposition 1 to derive $\epsilon$-optimal contracts even in the multiple-customer setting with an inventory constraint (Section 4). We also show how to adjust our contracts and propose a more practical contract in Section 3.5 where the firm is not allowed to use arbitrarily high commissions.

![Figure 1](image-url)
3.4. Discussion of first best achievability and model primitives

We pause to put some of our model primitives into context and shed light into how they play into our \( \epsilon \)-optimality result. In order to do this, note that a key assumption in our single-customer model is that the customer valuation distribution is continuous. It is interesting to compare this to the case of a discrete valuation distribution, which we do next with a stylized example with two customer types (high and low).

Specifically, assume that the customer valuation takes value \( v_H \) with probability \( \zeta_H \) and \( v_L \) with probability \( \zeta_L \). We assume that \( v_H > v_L \) and \( \zeta_L + \zeta_H = 1 \). In this case, the firm’s contract consists of two values \( t_H \) and \( t_L \) corresponding to the transfers for sales at \( v_H \) and \( v_L \) respectively. Note that by Lemma 2 we can assume that there is no transfer in the absence of a sale. We next characterize the optimal contract.

**Lemma 3.** The optimal contract to the agency problem (SC) with the binary distribution is defined as follows

\[
\begin{align*}
t^*_L &= \frac{\psi}{\zeta_L}, \\
t^*_H &= \frac{\psi}{\zeta_H \zeta_L}.
\end{align*}
\]

Moreover the expected payment is given by \( \mathbb{E}[t] = \psi + \frac{\psi}{\zeta_L} \).

In the case of this binary distribution, we observe that the optimal (second-best) contract incurs a positive rent \( \psi / \zeta_L \) due to limited liability. In this case, the second-best solution can also be implemented by giving an upfront payment of \( \psi \) to the agent and a bonus of \( \psi / \zeta_L \) for every sale at \( v_H \).

We shed some light on what breaks down in this example in terms of first-best achievability. With a continuous distribution of \( V \), the firm can specify very large contract payments for extreme sale prices, which are also extremely unlikely since they can only happen if \( V \) is appropriately large. Thus, by balancing the probability of these and magnitude of reward in case they happen, the firm can make sure it does not increase its expected payment too much. In other words, the family of contracts from (3) can be interpreted as a sequence of “lottery”-like contracts where paying the agent little for achieving typical sale prices is balanced with extreme payouts for sale prices at the right tail of the valuation distribution. When the distribution is discrete, the probability of the highest possible sale price is bounded from below (by \( \zeta_H \) in our case), and thus the firm cannot arbitrarily increase the payout at this price without leading to non-vanishing rent.

Another less obvious structural assumption we use in our baseline model is that the action space is tied one-to-one to the valuation space and thus also continuous. This is crucial: one could imagine
another model where \( V \) were continuous but the firm restricted the price discretion of the agent to a discrete menu of prices (or to a bounded interval of prices not containing the full right tail of \( V \)). Our \( \epsilon \)-optimality result would break in such a model due to the restriction on prices. This result is somewhat counter-intuitive, since one could perhaps naively expect that the firm can benefit from providing the agent with some price guidance through some limits on the allowable prices. Our analysis shows that this is in fact counterproductive for the firm’s profit. Note that in such a scenario, there is a double effect: not only that the firm leaves some rent on the table for the agent, but it also reduces revenues since it decreases the feasible set of \( SC \).

3.5. The impact of price-bounded commissions on rent

Section 3.3 shows that in order to approach first best with limited liability, the contract requires arbitrarily large commissions. Indeed, as discussed in the previous section, the sequence of contracts that we use to approach first best converges to a lottery-type contract where the agent gets small payments most of the time and huge rewards in unlikely events. In this section, we show how to adapt the methodology we introduced Section 3.3 to obtain a more natural contract that is closer to typical contracts observed in practice.

Let us bound the contract’s payment by the following constraint:

\[
t(p) \leq p, \quad \forall p. \tag{p-bounded}
\]

This is a natural constraint that ensures that the firm does not reward more than the revenue generated on a per sale basis. A similar constraint was also used in Innes (1990) and interpreted as a limited liability constraint on the side of the firm. Note that in the lottery type contract obtained in Section 3.3, the extremely low payments for most sales were subsidizing the extremely high payments in case of a high revenue sale; in contrast, the \( (p\text{-bounded}) \) adjustment aims at avoiding this: bounding the overall magnitude of payments presumably requires to increase the typical payment of the contract to keep it participation feasible.

In addition, we remark that imposing this additional constraint has a natural and practically appealing interpretation. In particular, any contract satisfying \( (p\text{-bounded}) \) and \( (\text{limited liability}) \) can be written as \( t(p) = \alpha(p) \cdot p \) where \( 0 \leq \alpha(p) \leq 1 \) for all \( p \). Thus, such a contract can be interpreted as “dynamic revenue-sharing” contract, where the agent receives a sales price dependent percentage of the revenue. This contract mimics the well-known revenue-share contract (Cachon and Lariviere 2005, Carroll 2015) in which the principal shares a constant fraction \( 0 \leq \alpha \leq 1 \) of the revenue with the agent; with our notation, such a revenue-share contract can be written as

\[
t^\alpha(p) = \alpha \cdot p, \quad \forall p.
\]
In our setting, a constant (i.e., price independent) revenue-share contract $t^r(\cdot)$ is sub-optimal. Nevertheless, we next show that a slightly modified version of the contract introduced in the previous section is actually optimal even under $(p$-bounded). More precisely, consider the following contract, parametrized by a scalar $\alpha$

$$t^{pb}_\alpha(p) = \min \left\{ 1, \frac{\alpha}{p \cdot \bar{F}(p)} \right\} \cdot p, \quad \forall p. \quad (5)$$

This contract, similar in spirit to the previous family of contracts $t^M(\cdot)$, tries to keep $\bar{F}(p) \cdot t(p)$ constant. This time, instead of using a constant upper bound imposed by $(M$-bounded), the contract uses a linear function as an upper bound.

**Proposition 3.** There exists $\alpha^*$ such that a contract the contract $t^{pb}(\cdot) = t^{pb}_{\alpha^*}(\cdot)$ solves

$$\max_{t \in T} \{ \mathbb{E}[V - t(V)] : t(\cdot) \text{ satisfies (participation), (IC), (limited liability), (p$-bounded)} \}. \quad (5)$$

In the optimal contract, the firm shares a fraction $\alpha^{pb}(p) = \min\{1, \alpha^*/(p \cdot \bar{F}(p))\}$ of the revenue from every sale. Unlike the revenue share contract, this fraction is a function of the sale price. Under a monotone hazard rate property, for any constant $\alpha$, the function $\alpha/\bar{F}(p)$ is increasing convex in $p$ which implies that the functions $p$ and $\alpha/\bar{F}(p)$ intersect exactly twice. This implies that the commission kept by the firm drops to zero both for low and high prices and is strictly positive for intermediate prices. Figure 3 illustrates the shape of the contract and contrasts it to the other contracts introduced so far in the paper. We note that the constraint $(p$-bounded) prevents the firm from achieving first-best payoffs, which we illustrate numerically next.

**Numerical example.** We illustrate the best payoff achieved by the contract $t^{pb}(\cdot)$. In particular, we compute the rent $(\mathbb{E}[t^{pb}(V)] - \psi)/J^{FB}$ to understand what is the loss with respect to the first-best payoff. We also compute the rent of the best revenue-share contract to illustrate how our proposed contract improves upon this well-known type of contract. Note that both contracts depend on a
Table 1 Rent incurred by the constraint *(p-bounded)*. We report \((E[t] - \psi)/J^FB\) for various distributions and contracts. Note that we use \(\psi = 0.1\).

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Uniform</th>
<th>Exponential</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t^\alpha(\cdot))</td>
<td>25.50%</td>
<td>6.46%</td>
</tr>
<tr>
<td>(t^{pb}(\cdot))</td>
<td>7.06%</td>
<td>1.84%</td>
</tr>
</tbody>
</table>

single parameter \(\alpha\) and we can find \(t^{pb}(\cdot)\) as well as the best revenue-share contract \(t^\alpha(\cdot)\) using a one dimensional line search. For both types of contracts, note that when \(\alpha\) is too small, (IC) or (participation) may not be satisfied. Table 1 summarizes the results when the valuation distribution is either a uniform \([0,1]\) or an exponential with mean 1. We observe that imposing *(p-bounded)* induces a strictly positive rent that ranges from 1.84% to 7.06%. For both distributions, \(t^{pb}(\cdot)\) provides a significant improvement over a revenue-share contract.

There are two conclusions to be drawn from these experiments. First, the common revenue-share contract is not appropriate for achieving near first-best performance. On the positive however, \(t^{pb}(\cdot)\), which is effectively a dynamic revenue-share contract, can achieve most of the benefits on the contract from Section 3.3.

4. Multiple-customer and inventory limitations

In this section, we generalize our model and analysis to capture the fact that a typical firm might have an inventory of goods to sell and multiple customers to serve over time.

4.1. Multi-period price discretion, learning effort and inventory

As hinted above, in this generalized model there are multiple customers, each with unit demand, who arrive over time. Each discrete time period corresponds to a customer arrival; hence, in a slight abuse of terms, we use customers and time periods interchangeably. The firm is endowed with multiple units of inventory to sell. We denote the number of customers and time periods by \(N\), and the number of units of inventory the firm has to sell by \(K (\leq N)\). We assume that the customer valuations are independent and drawn according to a common distribution with density \(f(v)\) as in the single-customer model. For the multiple-customer model, we make the following assumption:

\[\text{Assumption 2. The valuation function has an increasing failure rate, i.e. } f(p)/\bar{F}(p) \text{ is non-decreasing in } p.\]

\[\text{This is a common assumption which is satisfied by several popular distributions, such as Gamma, Weibull, modified extreme value, and the truncated normal distribution (see, e.g., Lariviere 2006) We use it to guarantee uniqueness of an optimal uninformed price.}\]
Upon the arrival of customer $i$, the agent must make an effort decision $e_i \in \{0, 1\}$ to learn the WTP of customer $i$, which we denote by the random variable $V_i$. In case the agent exerts learning effort, that is $e_i = 1$, she may then use this extra information to set the best informed price $\theta_i(V_i)$. If the agent does not exert effort to learn, she chooses a constant price $\phi_i$. We let $Y_i$ be the sale revenue at time $i$, i.e. $Y_i$ is equal to the price $p_i$ set by the agent in case a sale happens and to 0 in the absence of a sale, $Y_i = p_i \cdot 1\{V_i \geq p_i\}$. Since effort is not observable, the contract must depend only on the entire sequence of observable firm sale revenues. As in the single-customer case, for any sequence $(Y_1, \ldots, Y_N)$, let $T(Y_1, \ldots, Y_N)$ be the transfer from the firm to the agent. We also refer to the function $T$ as the contract. For now, we do not make any specific assumptions on the structure of $T$ and note that in principle, the firm could wish to implement payment rules that are non-linear in the history of observed sale revenues.

The resulting problem is challenging due to the fact that the common inventory of goods links together the learning and pricing decisions over all customers. In other words, when a new customer arrives, the agent must not only account for the outcome of that potential sale, but also for its future impact (due to inventory depletion) on the sales that could be achieved from subsequent customers. This leads to a complex agent policy space, which is very difficult to understand from an optimization perspective. Additionally, such a model can lead to very complex and hard to implement optimal policies.

Given these difficulties, we study a fluid model which captures the main tradeoffs arising from the presence of multiple-customer and inventory. More precisely, both the firm’s and agent’s problem are assumed to be deterministic with continuous dynamics, and are obtained by replacing the discrete stochastic demand process by its rate, which evolves as a continuous process. Similar fluid approaches have often been employed to tractably analyze other dynamic problems in the revenue management literature (Gallego and Van Ryzin 1997, Maglaras and Meissner 2006). Although we do not prove such a result in this work, fluid models have been shown to be limits of their discrete counterparts under a law-of-large-numbers type of scaling where the potential demand and the capacity grow proportionally large (Gallego and Van Ryzin 1997).

Given that we are working with a fluid model, without loss of generality we allow the agent to use a randomized strategy where she decides to exert learning effort on customer $i$, i.e. set $e_i = 1$, with probability $q_i$. Thus, a policy $\pi = (q_i, \theta_i(\cdot), \phi_i)_{i=1}^N$ consists, for each customer $i$, of a probability of exerting effort $q_i \in [0, 1]$ and a pricing decision for each alternative: $\phi_i \geq 0$ denotes the uninformed price when not learning whereas $\theta_i(\cdot) > 0$ denotes the informed price policy used when learning. Let $\Pi$ be the set of policy defined as

$$\Pi = \{(q_i, \theta_i(\cdot), \phi_i)_{i=1}^N : q_i \in [0, 1], \theta_i(\cdot) > 0, \phi_i > 0, \forall i = \{1, \ldots, N\}\}.$$
For each policy $\pi \in \Pi$, let $I(\pi)$ denote its expected inventory consumption. More precisely,

$$I(\pi) = \sum_{i=1}^{N} q_i \cdot \mathbb{E}[1\{V_i \geq \theta_i(V_i)\}] + (1-q_i) \cdot \bar{F}(\phi_i).$$

The inventory constraint that we impose can be written as follows:

$$I(\pi) \leq K. \tag{inventory}$$

For a particular policy $\pi \in \Pi$, let $\text{Rev}(\pi)$ be the firm’s expected revenue for a particular agent policy $\pi \in \Pi$. We can write

$$\text{Rev}(\pi) = \sum_{i=1}^{N} q_i \cdot \mathbb{E}[\theta_i(V_i) \cdot 1\{V_i \geq \theta_i(V_i)\}] + \sum_{i=1}^{N} (1-q_i) \cdot \phi_i \cdot \bar{F}(\phi_i).$$

Finally, the agent’s utility for a particular contract transfer $T$ and policy $\pi \in \Pi$ can be written as

$$U(\pi, T) = T - C(\pi),$$

where $C(\pi)$ is the expected learning cost and can be written as

$$C(\pi) = \sum_{i=1}^{N} q_i \cdot \psi.$$

### 4.2. Incentive feasible contracts

In the single-customer setting, we focused on the case where the firm induced the agent to learn. In the multiple-customer setting, the space of policies that the firm may want to induce is more complex. In particular, we show in Section 4.3, that the firm may sometimes want the agent to learn on a fraction of the customers only. We denote by $\pi^* = (q_i^*, \theta_i^*(\cdot), \phi_i^*)_{i=1}^{N}$ the policy that the firm is trying to induce. The following incentive compatibility constraint must therefore be satisfied:

$$\pi^* \in \arg \max_{\pi \in \Pi} \{U(\pi, T) : \pi \text{ satisfies (inventory)}\}. \tag{IC-multi}$$

In the following, we say that $\pi^*$ is a *best-response* to $T$ if $(\pi^*, T)$ satisfies (IC-multi). Note that here we assume that the inventory constraint is part of the agent problem; practically, this amounts to an agent who knows and manages the inventory level. Similar to the single-customer case, we also have a participation constraint stating that the agent should always get a positive expected utility. More precisely, we have

$$\mathbb{E}[T - C(\pi)] \geq 0. \tag{participation-multi}$$
Putting everything together, the multiple-customer firm’s problem (MC) can be written as
\[
\max_{\pi^*} \mathbb{E}[\text{Rev}(\pi^*) - T]
\]
\[
\text{s.t. } \pi^* \in \arg \max_{\pi \in \Pi} \{ U(\pi, T) : \pi \text{ satisfies (inventory)} \}
\]
\[
\mathbb{E}[T - C(\pi)] \geq 0
\]
\[
T \geq 0.
\]
\[(\text{MC})\]

Note that, as in the single-customer setting, we also impose a limited liability constraint stating that the transfer \(T\) must be non-negative. More precisely, for every sequence of realized revenues \((y_1, \ldots, y_N)\), we impose that \(T(y_1, \ldots, y_N) \geq 0\).

4.3. First-best payoff

In this section, we formulate the new first-best problem that arises in the case of multiple-customer and limited inventory. For this problem, we show that there are three types of first-best policies that the firm may want to induce: learn for all customers (“full learning”), for no customers (“no learning”), and a third policy where learning happens for only a fraction of the customers (“partial learning”). As we shall see, these regimes are closely related to a quantity that describes the firm’s opportunity cost of inventory.

Similar to the single-customer case, our benchmark is the problem where effort is observable and verifiable and can be included in the contract. We refer to this benchmark as first best. It can succinctly be written as
\[
J_{FB}^{MC} = \max_{\pi \in \Pi} \{ \mathbb{E}[\text{Rev}(\pi) - T] : (\pi, T) \text{ satisfy (inventory), (participation-multi)} \}.
\]

Similar to the single-customer case, for \(\epsilon > 0\), we say that a contract \(T\) is \(\epsilon\)-optimal if \(\mathbb{E}[\text{Rev}(\pi) - T] \geq J_{FB}^{MC} - \epsilon\), where \(\pi\) is a best-response to the contract \(T\). In the following, we assume without loss of generality that the inventory constraint is binding, i.e. \(I(\pi) = K\). Indeed, if the inventory is not binding, then the problem decouples over time and it is optimal to repeatedly use the optimal contract for the single-customer problem. Thus, we focus on the more interesting case that in the first-best outcome, inventory is binding. As in the single-customer problem, the participation constraint is binding in any optimal solution, and thus the first-best payoff can be obtained by solving the following optimization problem where the expected payment \(\mathbb{E}[T]\) has been substituted by \(C(\pi) = \sum q_i \cdot \psi\) using (participation-multi):
\[
J_{FB}^{MC} = \max \sum_{i=1}^{N} q_i \cdot \mathbb{E}[\theta_i(V_i) \cdot 1\{V_i \geq \theta_i(V_i)\}] + \sum_{i=1}^{N} (1 - q_i) \cdot \phi_i \cdot \bar{F}(\phi_i) - \sum_{i=1}^{N} q_i \cdot \psi
\]
\[
\text{s.t. } \sum_{i=1}^{N} q_i \cdot \mathbb{E}[1\{V_i \geq \theta_i(V_i)\}] + \sum_{i=1}^{N} (1 - q_i) \cdot \bar{F}(\phi_i) = K
\]
\[
q_i \in [0, 1], \quad \theta_i(\cdot) \geq 0, \quad \phi_i \geq 0.
\]
\[(6)\]
In the following, we showcase several important properties of (6). Intuition dictates that when there are more customers than inventory, the firm faces an opportunity cost of using a unit of inventory to make a sale. We first formalize this notion by relaxing the inventory constraint in (6) and solving for its optimal Lagrangian multiplier $\lambda$ which relates to this opportunity cost. This is a crucial parameter for our analysis for two reasons:

1. We identify three types of optimal policies for the firm that directly depend on $\lambda$ and the ratio $\rho = K/N$, which represents a natural and intuitive measure of inventory scarcity. Two of the optimal policies are natural extensions of the optimal policy in the single-customer case, i.e. the firm can decide to either induce effort to learn the valuation of all arriving customers, or not induce effort on any customer. However, we show that the multiple-customer model also requires a third policy, which we call partial learning, where the firm only exerts effort on a fraction of the customers. In the subsequent, we use the superscripts $f$-l (full-learning), $n$-l (no-learning) and $p$-l (partial-learning) for these three policies respectively.

2. The multiplier $\lambda$ can be interpreted as a reservation price that a customer’s valuation competes against for a unit of inventory. We indeed show how this reservation price can be used to modify the contract family from Section 3.2 to yield a contract which is $\epsilon$-optimal in this generalized model.

Consider the following Lagrangian relaxation of (6) for $\lambda \geq 0$:

$$J_{\lambda}^{FB-MC} = \max_{\lambda \geq 0} \sum_{i=1}^{N} q_i \cdot (E[(\theta_i(V_i) - \lambda) \cdot 1\{V_i \geq \theta_i(V_i)\}] - \psi) + (1 - q_i) \cdot (\phi_i - \lambda) \cdot \bar{F}(\phi_i) + K \cdot \lambda$$

s.t. $q_i \in [0, 1]$, $\theta_i(\cdot) \geq 0$, $\phi_i \geq 0$.  

\hspace{10cm} (7)

Since any solution to (6) is a feasible solution to (7), we have for all $\lambda \geq 0$, $J_{\lambda}^{FB-MC} \leq J_{\lambda}^{FB-MC}$. We therefore naturally consider the following dual problem:

$$\min_{\lambda \geq 0} J_{\lambda}^{FB-MC}.$$  

\hspace{10cm} (8)

Before characterizing the optimal solution to (8), notice that the objective in (7) separates over each term of the sum. Moreover, for each $i$, if $V_i - \lambda \geq 0$, then the best informed price is equal to $V_i$, whereas if $V_i - \lambda < 0$ any sale leads to a loss so the best informed price is $\infty$; thus, in order to maximize \(E[(V_i - \lambda) \cdot 1\{V_i \geq \theta_i(V_i)\}]\), it is optimal to set

$$\theta_i(V_i) = V_i \cdot 1\{V_i \geq \lambda\} + \infty \cdot 1\{V_i < \lambda\}.$$  

In other words, when exerting learning effort, it is optimal to extract the value of a customer only if the value is greater than $\lambda$. The Lagrange multiplier naturally acts as a reservation price to
Atasu, Ciocan, Désir: Price delegation with learning agents

Consequently, we can rewrite $J_{FB}^{FB-MC}$ as follows. For any $\lambda \geq 0$,

$$J_{FB}^{FB-MC} = N \cdot \max \left\{ E[(V - \lambda)^+] - \psi, \max_{\phi > 0} \{(\phi - \lambda) \cdot \bar{F}(\phi)\} \right\} + K \cdot \lambda. \quad (9)$$

Under this new expression of $J_{FB}^{FB-MC}$, it is clear that (8) is a convex optimization problem. Recall that in the single-customer setting, the firm’s first-best payoff is obtained by comparing the expected payoff if the firm induces learning, $E[V] - \psi$, and the expected payoff if the firm does not induce any learning, $\max_{\phi} \{(\phi - \lambda) \cdot \bar{F}(\phi)\}$. In this Lagrangian relaxation formulation, the firm’s expected payoff is penalized. If the firm induces learning, the expected payoff becomes $E[(V - \lambda)^+] - \psi$. In the absence of learning, the expected payoff becomes $\max_{\phi > 0} \{(\phi - \lambda) \cdot \bar{F}(\phi)\}$. Note that the optimal multiplier $\lambda$ depends on the inventory level $K$ and the number of customers $N$. Even under Assumption 1, when this penalty is too high, the firm will prefer to forego learning. We formalize an important quantity characterizing the boundary between learning and not learning in the next lemma.

**Lemma 4.** For all $\lambda \geq 0$, let $\Delta(\lambda) = N \cdot (E[(V - \lambda)^+] - \psi) - N \cdot \max_{\phi > 0} \{(\phi - \lambda) \cdot \bar{F}(\phi)\}$. There exists a unique $\lambda^{p-l}$ such that $\Delta(\lambda^{p-l}) = 0$. Moreover, we have $\Delta(\lambda) \geq 0$ if and only if $\lambda \leq \lambda^{p-l}$.

Intuitively, under the multiplier $\lambda^{p-l}$, the firm is indifferent between inducing learning and not inducing learning. Additionally, we define $\lambda^{f-l}$ and $\lambda^{n-l}$ as follows

$$\lambda^{f-l} = \arg \min_{\lambda \geq 0} \left\{ N \cdot (E[(V - \lambda)^+] - \psi) + K \cdot \lambda \right\}$$

$$\lambda^{n-l} = \arg \min_{\lambda \geq 0} \left\{ N \cdot \max_{\phi > 0} \{(\phi - \lambda) \cdot \bar{F}(\phi)\} + K \cdot \lambda \right\}.$$  

Both $\lambda^{f-l}$ and $\lambda^{n-l}$ are uniquely defined since the two functions they are respectively minimizing are convex and coercive. Moreover, we show in the proof of Lemma 5 that $\bar{F}(\lambda^{f-l}) = \rho$ and $\bar{F}(\phi^*(\lambda^{n-l})) = \rho$, where for all $\lambda$, we let $\phi^*(\lambda)$ denote the maximizer of $\{(\phi - \lambda) \cdot \bar{F}(\phi)\}$ which is unique by Assumption 2. Both of these quantities are defined so as to satisfy inventory feasibility conditions. For example, if the firm induces learning on all customers, inventory consumption will be equal to $N \bar{F}(\lambda) = K$; thus $\lambda^{f-l}$ is the multiplier for which inventory is fully consumed when the firm induces learning. Similarly, for $\lambda^{n-l}$ inventory is fully consumed in the absence of learning.

In the following, we will show these three multipliers define three distinct regimes for the firm’s first best policy: one where inducing learning dominates and all customers receive effort, an opposite
regime where no customers receive effort, and a regime where the firm faces a tie between inducing learning and not inducing learning. This is resolved by inducing learning only on a fraction of customers.

In order to do this, we first show that the firm’s optimal Lagrange multiplier $\lambda$ can take only one of the three discrete values defined above.

**Lemma 5.** Let $\lambda^*$ be the optimal solution to (8). We have

$$
\lambda^* = \begin{cases} 
    \lambda^{n-l}, & \text{if } \rho \leq \bar{F}(\phi^*(\lambda^{p-l})), \\
    \lambda^{p-l}, & \text{if } \bar{F}(\phi^*(\lambda^{p-l})) \leq \rho \leq \bar{F}(\lambda^{p-l}), \\
    \lambda^{f-l}, & \text{if } \rho \geq \bar{F}(\lambda^{p-l}).
\end{cases}
$$

We can now formally describe the regimes hinted at in the above. Since $\lambda^{p-l}$ does not depend on $\rho$, Lemma 5 indicates that there are three regimes depending on how the inventory scarcity parameter $\rho$ relates to the thresholds $\bar{F}(\lambda^{p-l})$ and $\bar{F}(\phi^*(\lambda^{p-l}))$. We naturally define three corresponding policies for the firm. The first one, $\pi^{f-l}$, is to induce learning on all customers and is formally defined as follows

$$
\pi^{f-l} = (1, V_i \cdot 1\{V_i \geq \lambda^{f-l}\} + \infty \cdot 1\{V_i < \lambda^{f-l}\}, \phi_i)_{i=1}^N.
$$

Under this policy, the firm induces the agent to learn the valuation $V_i$ of each customer and price exactly at $V_i$ if $V_i \geq \lambda^{f-l}$. Note that since $q_i = 1$, any value of the uninformed price $\phi_i$ leads to the same effective policy in this case and thus we can choose this arbitrarily. The opposite policy is to never learn the valuation of any customer and use an uninformed price $\phi^*(\lambda^{n-l})$. We denote this policy by $\pi^{n-l}$, defined as

$$
\pi^{n-l} = (0, \theta_i(\cdot), \phi^*(\lambda^{n-l}))_{i=1}^N.
$$

Note that since $q_i = 0$, any value of $\theta_i(\cdot)$ leads to the same policy in this case. For the third regime, we define a hybrid policy which learns on a fraction $\eta$ of the customers where $\eta$ is chosen such that

$$
\eta \bar{F}(\lambda^{p-l}) + (1 - \eta) \bar{F}(\phi^*(\lambda^{p-l})) = \rho,
$$

i.e. such that the policy is inventory feasible. We denote this policy by $\pi^{p-l}$ and formally define it as

$$
\pi^{p-l}(\eta, V_i \cdot 1\{V_i \geq \lambda^{p-l}\} + \infty \cdot 1\{V_i < \lambda^{p-l}\}, \phi^*(\lambda^{p-l}))_{i=1}^N.
$$

The next proposition relates these policies to the different regimes of $\lambda$ from Lemma 5. More precisely, we show that one of the three policies that we defined is always optimal.
Proposition 4. The optimal solution to (6) is

\[
\pi^* = \begin{cases} 
\pi^{n-1}, & \text{if } \rho \leq \bar{F}(\phi^*(\lambda^{p-1})), \\
\pi^{p-1}, & \text{if } \bar{F}(\phi^*(\lambda^{p-1})) \leq \rho \leq \bar{F}(\lambda^{p-1}), \\
\pi^{f-1}, & \text{if } \rho \geq \bar{F}(\lambda^{f-1}). 
\end{cases}
\]

Proposition 4 gives conditions under which the firm benefits from inducing learning. In particular, when the inventory is plenty, i.e. \(\rho \geq \rho^+ = \bar{F}(\lambda^{p-1})\), then it is optimal for the firm to induce learning of the valuation of every customer and extract all the value of each customer whose value is above the reservation price \(\lambda^{f-1}\), which is chosen so that the inventory is completely depleted in expectation. When the inventory is low, i.e. \(\rho \leq \rho^- = \bar{F}(\phi^*(\lambda^{p-1}))\), then the firm is better off never inducing learning and prefers the uninformed price \(\phi^*(\lambda^{f-1})\). For intermediate values of \(\rho\), the firm’s optimal strategy is to induce learning only on a fraction of the customers. Figure 3 illustrates the different first-best policies as a function of the inventory scarcity parameter \(\rho = K/N\).

We end this section by remarking that our Lagrangian relaxation is indeed tight. This is a direct corollary of Proposition 4, where we show that the solution \(\pi^*\) is optimal since it is feasible and yields an objective function equal to the upper bound \(J_{FB-MC}^{FB-MC}\).

Corollary 4. \(J_{FB-MC}^{FB-MC} = J_{\lambda^*}^{FB-MC} = \min_{\lambda \geq 0} J_{\lambda}^{FB-MC}\).
In the next section we show that this first-best solution can be implemented up to $\epsilon$-optimality even when the effort is not observable. As we shall see, this implementation will hinge on interpreting $\lambda^*$ as a reservation price which the firm can use to throttle the agent’s inventory consumption and guarantee inventory feasibility.

4.4. $\epsilon$-optimal contracts

We begin this section by remarking that one might expect an optimal or near-optimal multiple-customer contract $T$ to have a significantly more complex structure than a single-customer contract due to the need to coordinate how inventory is depleted across multiple sale opportunities. Indeed, when adding dynamics to principal-agent models, optimal contracts have been found to depend on the entire sequence sequence of sales $Y_1, \ldots, Y_N$ achieved by the agent in non-linear and non-separable fashions (see for example Rogerson 1985, for such a contract). Although the fluid simplification we adopt should make the contract independent of remaining inventory, there is a priori no reason to expect it should simplify the contract in any further way. For example, the optimal $T$ could have a bonus (or more generally convex) structure, where the bonus depends on the total sale revenues the agent achieves at the end of the time horizon via a threshold, e.g. $T(Y_1, \ldots, Y_N) = b \cdot 1\{\sum Y_i \geq \tau\}$.

In the sequel, we showcase a contract that enables the firm to still achieve $\epsilon$-optimality for the multiple-customer problem (MC). What is surprising is that we are able to achieve this with a contract which, in light of the previous discussion, may appear to be too simplistic: specifically, the contract is a separable contract of the form $T(Y_1, \ldots, Y_N) = \sum_{i=1}^N t_i(Y_i)$, treating each customer sale opportunity as independent from the other ones, and awarding a payment upon each individual customer sale. Moreover, the payment rule is static over customers, i.e. $t_i = t \in \mathcal{T}$.

Our strategy for constructing such a contract is as follows. For every $t \in \mathcal{T}$, we denote by $T = \{t\}$ the multiple-customer contract induced by the static single-customer contract $t$. In other words, $T$ depends separably on the sale revenue generated by the $i$-th customer, i.e. for any $(Y_1, \ldots, Y_N)$, we have

$$T(Y_1, \ldots, Y_N) = \sum_{i=1}^N t(Y_i).$$

Additionally, mimicking our contract from the single-customer setting, we assume that $t \in \mathcal{T}$ is non-decreasing, $t(\cdot) \geq 0$ and $t(0) = 0$. Note that this contract satisfies a stronger notion of limited liability since we impose each transfer $t_i$ to be non-negative as opposed to the aggregate transfer $T$. Also, note that in our model the transfer $T$ is realized after all sales occur; however, if $T$ is separable as above, it can be broken into individual transfers tied to each customer, each happening at the end of time period $i$. 
A potential issue with this parametric form for \( T \) is that it seems to ignore the inter-dependence due to limited inventory between the multiple customers. Our construction will effectively show that the firm can encode this inter-dependence through embedding a reservation price into the single-customer contract, which serves to throttle whether a given customer should be allocated inventory. Having built this reservation price on top of the \( \epsilon \)-optimal payment for the single-customer problem, the resulting multiple-customer will also inherit the first best approachability properties from the baseline model.

As a prequel to this result, we start by revisiting the single-customer case in the presence of the reservation price hinted to in the previous paragraph. Consider adding a reservation price to the single-customer setting. In particular, for some \( \lambda > 0 \), consider the following constraint:

\[ t(p) = 0, \quad \forall p \leq \lambda. \quad \text{(reservation)} \]

We show that a variant of our \( M \)-bounded contracts defined in Section 3.2 is optimal with this additional constraint. More precisely, for all \( M > 0 \) and \( \mu \geq \lambda \), consider the following contract:

\[
t^{M}_{\lambda, \mu}(p) = \begin{cases} 
0, & \text{if } p < \lambda, \\
M \cdot \frac{\bar{F}(\mu)}{\bar{F}(p)}, & \text{if } \lambda \leq p \leq \mu, \\
M, & \text{if } p > \mu,
\end{cases} \tag{10}
\]

We can once again leverage the machinery developed for Proposition 1 to prove that a natural modification of the \( t^{M}(\cdot) \) family of contracts suffices to solve the agency problem in the presence of (reservation). In this case, we can show the following.

**Lemma 6.** There exists \( \bar{M} \) such that for every \( M \geq \bar{M} \), there exists a threshold \( \mu = \mu(\lambda, M) \) such that \( t^{M}_{\lambda, \mu(\lambda, M)}(\cdot) \) solves

\[
\max_{t \in T} \{ \mathbb{E}[V - t(V)] : t(V) \text{ satisfies (participation), (IC), (limited liability), (M-bounded), (reservation)} \}.
\]

We now return to the multiple-customer problem and show that the multiple-customer contract induced by \( t^{M}_{\lambda, \mu(\lambda, M)}(\cdot) \) is \( \epsilon \)-optimal. As hinted, at a high level, the constraint (reservation) enforced directly on the contract payoff incentivizes the agent to follow the firm’s desired inventory feasible policy. Lastly, note that this contract can be made to depend on initial inventory level \( K \) and number of customers \( N \) through specifically adjusting the parameters \( \lambda, \nu \) and \( M \). However, it does not depend dynamically on the inventory trajectory. Also, recall that by Lemma 2, we can always set \( t(0) = 0 \) when imposing limited liability.

Recall that under a contract \( T = \{t\} \) for \( t \in T \), for each customer \( i \), the agent can decide to:
1. Not exert any effort to learn \((e_i = 0)\), price according to some price \(\phi_i\) and get a payment of \(t(\phi_i)\) when a sale happens.

2. Learn the willingness-to-pay \(V_i\) by exerting learning effort \((e_i = 1)\), bear the cost \(\psi_i\), choose some price function \(\theta_i\) and then get a payment of \(t(\theta_i(V_i))\) when a sale happens.

Thus, under a contract \(T = \{t\}\) and policy \(\pi \in \Pi\), we can write the expected payment of the firm to the agent \(P(\pi, T) = \mathbb{E}[T]\) as:

\[
P(\pi, T) = \sum_{i=1}^{N} q_i \cdot \mathbb{E}[t(\theta_i(V_i)) \cdot 1\{V_i \geq \theta_i(V_i)\}] + (1 - q_i) \cdot t(\phi_i) \cdot \bar{F}(\phi_i)
\]

In order to show that a contract \(T\) achieves \(\epsilon\)-optimality, it is sufficient to show that \(P(\pi, T)\) tends to \(\psi\), if we set \(\pi\) to be the agent’s best-response to \(T\), i.e. if \((\pi, T)\) satisfies \((\text{lC-multi})\).

Instead of solving \((\text{MC})\) directly, we show that the reservation contracts of the type \((10)\) implement the policy \(\pi^*\) from Proposition 4 and are \(\epsilon\)-optimal. Note that if the firm wants the agent to implement a particular target policy \(\pi^*\), it must find a contract \(T = \{t\}\) such that \((\pi^*, T)\) satisfies \((\text{lC-multi})\) and \((\text{participation-multi})\).

In fact, we begin by showing in the next lemma that \(T = \{t_{M,\lambda,\mu}(\cdot)\}\) implements \(\pi_{f}^{\text{f-l}}\) or \(\pi_{p}^{\text{p-l}}\) for carefully chosen sequences of the parameters \(M, \lambda, \mu\). For every \(M > 0\), recall that the contract is defined by a reservation price \(\lambda\) below which the agent does not get anything, and a threshold \(\mu\) above which the agent commission is capped at \(M\). To implement \(\pi_{f}^{\text{f-l}}\) or \(\pi_{p}^{\text{p-l}}\), we use the value of \(\lambda^*\) defined in Lemma 5 as a reservation price and then construct a threshold \(\mu\) which depends on \(\lambda\) and \(M\). We remark that it is not clear why using \(\lambda^*\) as a reservation value should be sufficient to implement either \(\pi_{f}^{\text{f-l}}\) or \(\pi_{p}^{\text{p-l}}\). In particular, the contract should enforce the feasibility of \((\text{inventory})\) in the agent’s problem; however it is not clear why \(\lambda^*\), which is the correct shadow price of inventory in the firm’s centralized problem, should be an optimal multiplier in the agent’s problem. We show in the next lemma that using our modified contract defined in \((10)\) induces the agent to implement first-best policies.

**Lemma 7.** For every \(M > \bar{M}\), there exist \(\mu_{f}^{\text{f-l}}\) and \(\mu_{p}^{\text{p-l}}\) such that

\[
(\pi_{f}^{\text{f-l}}, T_{f}^{\text{f-l}}) \text{ satisfies } (\text{lC-multi}) \text{ and } (\text{participation-multi}),
\]

\[
(\pi_{p}^{\text{p-l}}, T_{p}^{\text{p-l}}) \text{ satisfies } (\text{lC-multi}) \text{ and } (\text{participation-multi}),
\]

where \(T_{f}^{\text{f-l}} = \{t_{M,\lambda_{f}^{\text{f-l}},\mu_{f}^{\text{f-l}}}\}\) and \(T_{p}^{\text{p-l}} = \{t_{M,\lambda_{p}^{\text{p-l}},\mu_{p}^{\text{p-l}}}\}\).
Lemma 7 shows that $T^{f_l}$ and $T^{p_l}$ induce the agent to follow the first-best policies $\pi^{f_l}$ and $\pi^{p_l}$, respectively. We next show that the cost of implementing the first-best policy vanishes when $M \to \infty$. More specifically, the participation constraint \text{(participation-multi)} is binding asymptotically when $M \to \infty$.

**Lemma 8.** \( \lim_{M \to \infty} P(\pi^{f_l}, T^{f_l}) = \lim_{M \to \infty} P(\pi^{p_l}, T^{p_l}) = \psi. \)

This directly implies that the first-best payoff is achievable up to an additive factor of $\epsilon$ by a separable contract. In particular, we have the following result which generalizes Corollary 3 for the single-customer setting.

**Corollary 5.** Let $\pi^*$ be the optimal first-best policy. For every $\epsilon > 0$, there exists a separable contract $T = \{t\}$ which is $\epsilon$-optimal, i.e. $E[\text{Rev}(\pi^*) - T] \geq J_{FB-MC} - \epsilon$, and additionally $(\pi^*, T)$ satisfies \text{(participation-multi)} and \text{(IC-multi)}.

We pause to make a few observations about the above result and how it relates to our model. First, recall that in our model the agent directly enforces the inventory constraint. Although the inventory constraint does not appear explicitly in the corollary statement, it appears indirectly through \text{(IC-multi)}. An interesting observation arising from Lemma 7 and Corollary 5 is that, if the firm wanted to induce $\pi^{f_l}$, it could do so via $T^{f_l}$ whether or not the agent enforced the inventory constraint themselves, making our finding robust to whether the firm shares inventory information with the agent. This is because the zeroing of the contract below $\lambda = \lambda^{f_l}$ implicitly results in the agent not selling to any customer with valuation below $\lambda^{f_l}$, which automatically guarantees inventory feasibility. On the other hand, the contract shape by itself is not sufficient to guarantee feasibility if the firm prefers inducing $\pi^{p_l}$. There, the agent’s knowledge of $K$ is crucial in choosing the appropriate throttling parameter $\eta$ between learning and not learning in order to satisfy the inventory constraint. It is an interesting question to ask whether there are other contracts which do not require that the agent knows $K$ even in the partial learning case.

Another question is whether there is a unique mapping between the contracts $T^{f_l}$ and $T^{p_l}$ and the policies $\pi^{f_l}$ and $\pi^{p_l}$. To be more precise, one could ask for a stronger definition of implementability, stating for that a fixed $T$, $\pi^*$ is the unique best-response policy. We do not attempt this in this paper for simplicity, but doing so is possible. This would require adding a perturbation term to the above contracts to break any ties that could arise, leading to a more cumbersome analysis.
5. Conclusions and future research directions

In this paper, we embed a principal-agent problem into the classical revenue management setting of pricing with finite inventories. This model reflects the common practice of firms employing a salesperson to interact with customers and to make decentralized pricing decisions. We focus on the benefits that can be gained from the agent’s ability to exert effort towards extracting signals about customer valuations, which are then used to refine pricing. Overall, we consider our main contribution to be the study of a novel notion of incentive compatibility observed in the presence of pricing sales agents that can learn, along with a new technique (i.e., the primal-dual approach we utilize) producing $\epsilon$-optimal contracts in single-customer and multiple-customer with inventory constrained settings. We also believe that our results highlight the importance of revenue management research recognizing the role that managing pricing delegation and salesforce intermediation can play in practice.

We close by noting that there are other open research questions related to the problem we study in this paper. One question that we do not tackle is whether the $\epsilon$-optimality result for the multiple-customer problem transfers to a non-fluid, discrete version of the problem. As observed in Section 4, there is a large stream of literature in revenue management which shows that fluid approximation based solutions perform well under an appropriate scaling of customer volume and inventory. While a similar result may be possible here, the analysis can be quite challenging due to the interplay between between the firm’s and the agent’s respective optimization problems. In particular, while a policy that the firm desires to implement may be incentive compatible in the agent’s fluid problem, it would only be approximately incentive compatible in the agent’s non-fluid problem: defining and working with such a notion of approximate incentive compatibility is a non-trivial research problem by itself.

There are also several facets of our model that could be generalized. For example, we assume a binary learning effort. Examining continuous effort which translates into imperfect signals of customer valuations can be an interesting generalization. While we account for some form of risk aversion through enforcing limited liability, the agent’s utility function could be generalized to directly accommodate risk-aversion. Another avenue for research could be to investigate how the demand shifting sales effort the literature has often studied could interact with the learning effort we focus on, and whether these two play the roles of complements or substitutes.

Lastly, while salespeople are adept of acquiring personalized customer information to use for pricing, firms also have access to analytical technologies that allow them to aggregate information and optimize prices in order to support sales operations. To this end, combining centralized and
decentralized information in pricing decisions can be an effective approach, which has indeed been shown to substantially improve firm profits in Phillips et al. (2015)’s empirical study. To this end, designing salesforce contracts that can facilitate effectively embedding centralized technology based pricing recommendations in delegated pricing decisions of learning agents is a promising research direction.

References


Proofs of Statements

EC.1. Proofs of Section 3

EC.1.1. Proof of Corollary 1

By contradiction, assume that \((\text{participation})\) is binding, i.e. \(\mathbb{E}[t(V)] = \psi\). Using \((\text{IC})\), this implies that \(0 \geq \max_{\psi \geq 0} \{t(\psi) \cdot \bar{F}(\psi) + t(0) \cdot F(\psi)\}\) which in turns implies that \(t(p) = 0\) for all \(p\) by \((\text{limited liability})\). This is a contradiction which shows that limited liability and incentive compatibility cannot be both satisfied. \(\square\)

EC.1.2. Proof of Lemma 2

For all \(p\), let \(\tilde{t}(p) = t(p) - t(0)\). We first show that \(\tilde{t}(\cdot)\) is a feasible contract. Since \(t(\cdot)\) is non-decreasing by assumption, so is \(\tilde{t}(\cdot)\). This also implies that \(t(p) \geq t(0)\) for all \(p\) which immediately implies that \(\tilde{t}(\cdot)\) satisfies \((\text{limited liability})\). Next, for all \(p > 0\), we have

\[
\mathbb{E}[\tilde{t}(V) - \psi] = \mathbb{E}[t(V) - \psi] - t(0) \\
\geq t(p) \cdot \bar{F}(p) + t(0) \cdot F(p) - t(0) \\
\geq \tilde{t}(p) \cdot \bar{F}(p) + \tilde{t}(0) \cdot F(p).
\]

This shows that \(\tilde{t}(\cdot)\) satisfies \((\text{IC})\) as well as \((\text{participation})\). Finally, since \(\tilde{t}(p) \leq t(p)\) for all \(p\), we have \(\mathbb{E}[V - \tilde{t}(V)] \geq \mathbb{E}[V - t(V)]\). \(\square\)

EC.1.3. Proof of Corollary 2

We have for \(p \leq \mu\),

\[
\frac{\partial \tilde{t}(p)}{\partial p} = \frac{f(p)}{F(p)^2} \cdot M \cdot \bar{F}(\mu),
\]

which means that under an increasing hazard rate assumption, \(t(p)\) is convex increasing in \(p\) for \(p \leq \mu\). \(\square\)

EC.1.4. Proof of Proposition 1

We begin by showing that there exists \(M_1\) such that for all \(M \geq M_1\), (4) is feasible. In particular, we show that for \(M\) large enough, there always exists \(\mu\) such that \(t^M(\cdot)\) satisfies \((\text{IC})\). First, note that for all \(M\) and \(\mu\), we have

\[
\mathbb{E}[t^M_\mu(V)] = M \cdot \int_0^\mu \frac{\bar{F}(\mu)}{F(v)} \cdot f(v) \cdot dv + M \cdot \bar{F}(\mu) = M \cdot \bar{F}(\mu) \cdot [1 - \log(\bar{F}(\mu))].
\]
Moreover,
\[
\max_p \{t^M_\mu(p) \cdot \bar{F}(p)\} = M \cdot \bar{F}(\mu).
\]

Consequently, it suffices to show that there exists \( \mu \) such that
\[
M \cdot \bar{F}(\mu) \cdot [1 - \log(\bar{F}(\mu))] - \psi \geq M \cdot \bar{F}(\mu) \iff \bar{F}(\mu) \cdot \log(\bar{F}(\mu)) \geq \frac{\psi}{M}.
\]

Note that the function \(-\bar{F}(x) \cdot \log(\bar{F}(x))\) is upper bounded for \( x \in [0, 1] \). Consequently, for \( M \geq \psi / \max_{x \in [0, 1]} \{\bar{F}(x) \cdot \log(\bar{F}(x))\} \), there always exists \( \mu \) such that \( t^M_\mu(\cdot) \) satisfies (IC). We now assume that \( M \geq M_1 \) so that (4) is feasible and proceed to identifying an optimal solution.

In the objective function of the firm \( \mathbb{E}[V - t(V)] \), note that \( \mathbb{E}[V] \) is constant. Consequently, we can formulate (4) as follows

\[
\begin{align*}
z^*(M) &= \min_{z,t \in T} z \\
\text{s.t.} \quad &z \geq \psi + \max_p \{t(p) \cdot \bar{F}(p)\} \\
&z = \mathbb{E}[t(V)] \\
&0 \leq t(p) \leq M, \quad \forall p.
\end{align*}
\tag{EC.1}
\]

Since the problem is feasible, \( z^*(M) \) exists. Moreover, \( z \) is lower bounded by \( \psi \), and consequently we have \( z^*(M) > -\infty \). We show that if we fix \( z \) to its optimal value \( z^*(M) \), there exists an optimal contract of the form \( t^M_\mu(\cdot) \) for some \( \mu \). Given \( z^*(M) \), the optimal contract can be found by solving the following problem where the decision variable is \( t(p) \) only:

\[
\begin{align*}
\min_{t(p)} &\max_p \{t(p) \cdot \bar{F}(p)\} \\
\text{s.t.} \quad &z^*(M) = \mathbb{E}[t(V)] \\
&0 \leq t(p) \leq M, \quad \forall p.
\end{align*}
\]

Even though the decision variable is a function \( t(\cdot) \), the problem can be written as the following continuous linear program using the epigraph formulation:

\[
\begin{align*}
\min &\quad y \\
\text{s.t.} \quad &y \geq t(p) \cdot \bar{F}(p), \quad \forall p \\
&z^*(M) = \mathbb{E}[t(V)] \\
&0 \leq t(p) \leq M, \quad \forall p.
\end{align*}
\tag{primal}
\]

For all \( \nu(\cdot) \geq 0 \), \( \zeta(\cdot) \geq 0 \), and \( \xi(\cdot) \geq 0 \), the Lagrangian of (primal) is

\[
\mathcal{L}(t(\cdot), y, \nu(\cdot), \zeta, \xi(\cdot)) = y + \int (t(p) \cdot \bar{F}(p) - y) \cdot \nu(p) \cdot dp + \zeta \cdot \left( z^*(M) - \int t(p) \cdot f(p) \cdot dp \right)
\]
\[
+ \int (t(p) - M) \cdot \xi(p) \cdot dp
\]
\[
= \zeta \cdot z^*(M) - M \cdot \int \xi(p) \cdot dp + y \cdot \left(1 - \int \nu(p) \cdot dp\right)
\]
\[
+ \int t(p) \cdot (\bar{F}(p) \cdot \nu(p) - \zeta \cdot f(p) + \xi(p)) \cdot dp,
\]
where we have not dualized the non-negativity constraint on \( t(p) \). Minimizing over \( t(p) \geq 0 \), this leads to the following dual problem:

\[
\max \ \zeta \cdot z^*(M) - M \cdot \int \xi(p) \cdot dp
\]
\[
s.t. \ \int \nu(p) \cdot dp = 1, \quad (\text{dual})
\]
\[
\bar{F}(p) \cdot \nu(p) + \xi(p) \geq \zeta \cdot f(p), \forall p,
\]
\[
\nu(p) \geq 0, \xi(p) \geq 0.
\]

We now exhibit a pair of feasible primal/dual solutions for which the primal and dual objective functions are equal. This shows that these are optimal for their respective problem since weak duality holds for our problem (Reiland 1980).

1. **Primal solution.** Consider the following feasible primal solution defined as

\[
t(p) = \min \left\{ M \cdot \bar{F}(\mu^*), M \right\}, \quad \forall p,
\]

where \( \mu^* \) is chosen so that

\[
\int \min \left\{ M \cdot \bar{F}(\mu^*), M \right\} \cdot f(p) \cdot dp = z^*(M). \quad (\text{EC.2})
\]

We can rewrite the above as

\[
M \cdot \int \mu^* \frac{\bar{F}(\mu^*)}{\bar{F}(p)} \cdot f(p) \cdot dp + \bar{F}(\mu^*) \cdot M = z^*(M)
\]
\[
\Rightarrow M \cdot \bar{F}(\mu^*) \cdot \left( \int_0^{\mu^*} \frac{f(p)}{\bar{F}(p)} \cdot dp + 1 \right) = z^*(M)
\]
\[
\Rightarrow \bar{F}(\mu^*) \cdot (1 - \log(\bar{F}(\mu^*))) = \frac{z^*(M)}{M}. \quad (\text{EC.3})
\]

Note that the function \( \bar{F}(x) \cdot (1 - \log(\bar{F}(x))) \) is a decreasing function of \( x \). Moreover, \( \bar{F}(0) \cdot (1 - \log(\bar{F}(0))) = 1 \) and \( \lim_{x \to \infty} \bar{F}(x) \cdot (1 - \log(\bar{F}(x))) = 0 \). Consequently, if \( z^*(M) \leq M \), there exists a unique \( \mu^* \) satisfying (EC.2). For this choice of \( t(\cdot) \), the minimum value that \( y \) can take is \( y = M \cdot \bar{F}(\mu^*) \). Consequently, the corresponding objective function of (primal) can be rewritten as

\[
M \cdot \bar{F}(\mu^*) = \frac{z^*(M) - M \cdot \bar{F}(\mu^*)}{\log(\frac{1}{(\bar{F}(\mu^*)})}. \quad (\text{EC.3})
\]
2. Dual solution. We now consider the following feasible dual solution

\[ \nu(p) = \begin{cases} 
    \frac{f(p) \cdot C}{F(p)}, & \text{if } p \leq \mu^*, \\
    0, & \text{otherwise,}
\end{cases} \]

\[ \xi(p) = \begin{cases} 
    f(p) \cdot C, & \text{if } p \geq \mu^*, \\
    0, & \text{otherwise,}
\end{cases} \]

where \( C \) is chosen such that

\[ \int \nu(p) \cdot dp = 1 \implies C \cdot \left( \int_0^{\mu^*} \frac{f(p)}{F(p)} dp \right) = 1 \implies C \cdot \log \left( \frac{1}{F(\mu^*)} \right) = 1. \]

For this pair of functions, the choice of \( \zeta \) which maximizes the objective of (dual) is \( \zeta = C \). Consequently, the dual objective can be written as

\[ C \cdot z^*(M) - C \cdot \int_{p \geq \mu^*} M \cdot f(p) \cdot dp = C \cdot (z^*(M) - M \cdot \bar{F}(\mu^*)) = M \cdot \bar{F}(\mu^*), \]

where the last equation holds by (EC.3).

This shows that for this pair of primal/dual solutions, the primal and dual objective functions are equal and therefore these are optimal solutions. To conclude the proof, note that \( z^*(M) \) is a decreasing function of \( M \), therefore there exists \( \bar{M} \geq M_1 \) such that for all \( M \geq \bar{M} \), we have \( z^*(M) \leq M \).

We also comment on the fact that this proof yields an algorithm to solve problem (4). Specifically, note that for a fixed \( z \leq M \), we can replace \( z^*(M) \) with \( z \) in (primal) and obtain an infinite dimensional linear program in \( y \) and \( t(\cdot) \) which is tractable to solve; if the program is feasible, then \( z^*(M) \leq z \), whereas if it is infeasible, \( z^*(M) > z \). Thus, to find the true \( z^*(M) \), we can perform a grid search or a more sophisticated search procedure until \( z \) cannot be lowered while maintaining feasibility of the continuous linear program. At that point, \( z^*(M) = z \). \( \square \)

**EC.1.5. Proof of Proposition 2**

Let \( z^*(M) \) be the optimal solution defined in (EC.1). By definition of the limited liability rent, we have for all \( M \), \( z^*(M) = \psi + R(M) \). As \( M \) increases, the feasible set increases and therefore both \( z^*(M) \) and \( R(M) \) are decreasing functions of \( M \). Since they are both lower bounded, \( z^*(M) \geq \psi \) and \( R(M) \geq 0 \), they both converge when \( M \to \infty \). Using (EC.2), we have

\[ M \cdot \bar{F}(\mu(M)) \cdot \left( \int_0^{\mu(M)} \frac{f(p)}{F(p)} \cdot dp + 1 \right) = z^*(M) \]

\[ \implies R(M) \cdot \left( \log \left( \frac{1}{F(\mu(M))} \right) + 1 \right) = \psi + R(M) \]

\[ \implies R(M) = -\frac{\psi}{\log \left( F(\mu(M)) \right)}. \]
This implies that $\mu(M)$ is an increasing function of $M$. Additionally, we have

$$F(\mu(M)) \cdot (1 - \log(F(\mu(M)))) = \frac{z^*(M)}{M}.$$ 

Since $z^*(M) \geq \psi$, $z^*(M)/M \to 0$ when $M \to \infty$. Consequently, it must be that $F(\mu(M)) \to 0$ which in turns implies that $R(M) \to 0$. Note that $\mu(M)$ converges to $\infty$ if the support of the valuation distribution is unbounded and to the upper bound of the support otherwise. □

**EC.1.6. Proof of Lemma 3**

The firm’s problem in the presence of limited liability can be written as the following linear program.

$$
\begin{align*}
\max & \quad \zeta_H(v_H - t_H) + \zeta_L(v_L - t_L) \\
& \quad \zeta_H t_H + \zeta_L t_L - \psi \geq \max \{\zeta_H t_H, t_L\} \\ & \quad \zeta_H t_H + \zeta_L t_L - \psi \geq 0 \\
& \quad t_H, t_L \geq 0.
\end{align*}
$$

Note that if the participation constraint is binding, i.e. $\zeta_H t_H + \zeta_L t_L = \psi$, then the incentive constraint implies that $0 \geq \max \{\zeta_H t_H, t_L\}$ which implies that $t_H = t_L = 0$ by the limited liability constraint which is a contradiction. Consequently, the participation constraint is not binding and it is easy to see that the non-negativity payment constraints cannot be binding for an extreme point solution. Solving for the incentive constraints for $t_H$ and $t_L$, we get the following solution:

$$
\begin{align*}
\zeta_L t_L &= \psi, \\
\zeta_H t_H &= \psi + \zeta_H t_L = \psi + \frac{\zeta_H}{\zeta_L} \psi = \frac{1}{\zeta_L} \psi.
\end{align*}
$$

Under this contract, the expected payment is $\zeta_L t_L + \zeta_H t_H = \psi + \psi/\zeta_L$. This concludes the proof. □

**EC.1.7. Proof of Proposition 3**

By Assumption 1, (5) is feasible. Indeed, a linear contract of the form $t(p) = p$ for all $p$ provides a feasible solution to (5). The rest of the proof parallels that of Proposition 1. In particular, (5) can be reformulated as

$$
\begin{align*}
z^* = \min & \quad z \\
\text{s.t.} & \quad z \geq \psi + \max \{t(p) \cdot F(p)\} \\
& \quad z = E[t(V)] \\
& \quad 0 \leq t(p) \leq p, \quad \forall p.
\end{align*}
$$
Since \( z \) is lower bounded by \( \psi \), we have \( z^* > -\infty \). Given \( z^* \), the optimal contract can be found by solving the following problem where the decision variable is \( t(p) \) only:

\[
\begin{align*}
\min \ y \\
\text{s.t.} \quad t(p) \cdot \bar{F}(p) &\leq y, \quad \forall p \\
\end{align*}
\]

\[z^* = \mathbb{E}[t(V)]\]

\[0 \leq t(p) \leq p, \quad \forall p.\]

For all \( \nu(\cdot) \geq 0, \zeta \) and \( \xi(\cdot) \geq 0 \), the Lagrangian of this problem is

\[
\mathcal{L}(t(\cdot), y, \nu(\cdot), \zeta, \xi(\cdot)) = y + \int (t(p) \cdot \bar{F}(p) - y) \cdot \nu(p) \cdot dp + \zeta \cdot \left( z^* - \int t(p) \cdot f(p) \cdot dp \right) + \int (t(p) - p) \cdot \xi(p) \cdot dp
\]

\[= \zeta \cdot z^* - \int p \cdot \xi(p) \cdot dp + y \cdot \left( 1 - \int \nu(p) \cdot dp \right) + \int t(p) \cdot (\bar{F}(p) \cdot \nu(p) - \zeta \cdot f(p) + \xi(p)) \cdot dp,\]

where we have not dualized the non-negativity constraint on \( t(p) \). This leads to the following dual problem:

\[
\begin{align*}
\max \zeta \cdot z^* - \int p \cdot \xi(p) \cdot dp \\
\text{s.t.} \quad \int \nu(p) \cdot dp &= 1 \\
\bar{F}(p) \cdot \nu(p) + \xi(p) &\geq \zeta \cdot f(p), \quad \forall p \\
\mu(p) &\geq 0.
\end{align*}
\]

We now exhibit a pair of feasible primal/dual solutions for which the primal and dual objective functions are equal.

1. **Primal solution.** Consider the following feasible primal solution defined as

\[
t(p) = \min \left\{ p, \frac{\kappa}{\bar{F}(p)} \right\}, \quad \forall p,
\]

where \( \kappa \) is chosen so that

\[
\int \min \left\{ p, \frac{\kappa}{\bar{F}(p)} \right\} \cdot f(p) \cdot dp = z^*.
\]

The best choice of \( y \) for this given \( t(p) \) is \( y = \kappa \). We can rewrite the above equation as

\[
\int_{p: p \leq \kappa/\bar{F}(p)} p \cdot f(p) \cdot dp + \kappa \cdot \int_{p: p \geq \kappa/\bar{F}(p)} \frac{f(p)}{\bar{F}(p)} dp = z \implies \kappa = \frac{z^* - \alpha}{\beta}.
\]
where
\[
\alpha = \int_{p:p \leq \kappa/F(p)} p \cdot f(p) \cdot dp \\
\beta = \int_{p:p \geq \kappa/F(p)} f(p) \cdot \frac{1}{F(p)} dp.
\]

2. **Dual solution.** We now consider the following feasible dual solution
\[
\nu(p) = \begin{cases} 
    f(p) \cdot L \cdot \frac{1}{F(p)}, & \text{if } p \geq \kappa/F(p) \\
    0, & \text{otherwise}
\end{cases}
\]
\[
\xi(p) = \begin{cases} 
    f(p) \cdot L, & \text{if } p \leq \kappa/F(p) \\
    0, & \text{otherwise}
\end{cases}
\]

where \( L \) is chosen such that
\[
\int \nu(p) \cdot dp = 1 \implies L = \left( \int_{p:p \geq \kappa/F(p)} f(p) \cdot \frac{1}{F(p)} dp \right)^{-1} = \frac{1}{\beta}.
\]

For this pair of functions, the choice of \( \zeta \) which maximizes the dual objective is \( \zeta = L \). Consequently, the dual objective can be written as
\[
L \cdot z^* - L \cdot \int_{p:p \leq \kappa/F(p)} p \cdot f(p) \cdot dp = \frac{z^* - \alpha}{\beta} = \kappa.
\]

This shows that for this pair of primal/dual solutions, the primal and dual objective functions are equal and therefore these are optimal solutions which concludes the proof. \( \square \)

**EC.2. Proofs of Section 4**

**EC.2.1. Proof of Lemma 4**

\( \mathbb{E}[(V - \lambda)^+] - \psi \) is a strictly decreasing convex function which goes to \( -\psi \) when \( \lambda \) goes to \( \infty \). On the other hand, \( \max_{\phi} \{(\phi - \lambda)F(\phi)\} \) is a strictly decreasing convex function which goes to 0 when \( \lambda \) goes to \( \infty \). Moreover, by Assumption 1, we have \( \mathbb{E}[(V - \lambda)^+] - \psi \geq \max_{\phi} \{(\phi - \lambda)F(\phi)\} \) when \( \lambda = 0 \). Consequently, these two functions intersect exactly once which implies the desired result. \( \square \)

**EC.2.2. Proof of Lemma 5**

Inspecting (9), observe that we can write for all \( \lambda \geq 0 \),
\[
J_{FB-MC}^\lambda = \max\{f(\lambda), g(\lambda)\},
\]
where \( f(\lambda) = N \cdot (\mathbb{E}[(V - \lambda^+) - \psi] + K\lambda \) and \( g(\lambda) = N \cdot \max_{\phi} \{ (\phi - \lambda)\tilde{F}(\phi) \} + K\lambda \). Note that both \( f(\lambda) \) and \( g(\lambda) \) are convex functions and therefore, so is \( J^{\text{FB-MC}}_{\lambda} \). Moreover, note that for all \( \lambda \geq 0 \), \( \Delta(\lambda) = f(\lambda) - g(\lambda) \) where \( \Delta(\lambda) \) was defined in Lemma 4. Recall that

\[
\lambda^{\text{fl}} = \arg\min_{\lambda \geq 0} \left\{ N \cdot (\mathbb{E}[(V - \lambda^+) - \psi]) + K \cdot \lambda \right\},
\]

\[
\lambda^{\text{nl}} = \arg\min_{\lambda \geq 0} \left\{ N \cdot \max_{\phi > 0} \{ (\phi - \lambda) \cdot \tilde{F}(\phi) \} + K \cdot \lambda \right\}.
\]

We begin by showing that \( \tilde{F}(\lambda^{\text{fl}}) = \tilde{F}(\phi^*(\lambda^{\text{nl}})) = \rho \). Writing the first-order condition for \( f(\lambda) \), \( \lambda^{\text{fl}} \) satisfies \( \tilde{F}(\lambda^{\text{fl}}) = \rho \). We now derive the same property of \( \lambda^{\text{nl}} \). Recall that for all \( \lambda \), \( \phi^*(\lambda) \) denotes the maximizer of \( \{ (\phi - \lambda) \cdot \tilde{F}(\phi) \} \) which is unique by Assumption 2. Consequently, writing the first-order condition with respect to \( \phi \), \( \phi^*(\lambda) \) must satisfy

\[
\tilde{F}(\phi^*(\lambda)) - (\phi^*(\lambda) - \lambda) f(\phi^*(\lambda)) = 0 \implies \phi^*(\lambda) = \lambda + \frac{\tilde{F}(\phi^*(\lambda))}{f(\phi^*(\lambda))}.
\]

(EC.4)

Note that the above also implies that \( \phi^*(\lambda) \) is a non-decreasing function of \( \lambda \). Writing the first-order condition (with respect to \( \lambda \) this time), we obtain

\[
0 = N \cdot \left( \frac{\partial}{\partial \lambda} \phi^*(\lambda) - 1 \right) \cdot \tilde{F}(\phi^*(\lambda)) - N \cdot (\phi^*(\lambda) - \lambda) \cdot f(\phi^*(\lambda)) \cdot \frac{\partial}{\partial \lambda} \phi^*(\lambda) + K
\]

\[
= -N \cdot \tilde{F}(\phi^*(\lambda)) + K
\]

\[
\implies \tilde{F}(\phi^*(\lambda^{\text{nl}})) = \rho,
\]

where we have simplified the first equality using (EC.4).

We now analyze the minimizer of \( J^{\text{FB-MC}}_{\lambda} \). Since we are minimizing the maximum of two convex functions, we have three cases to consider.

1. If \( f(\lambda^{\text{fl}}) \geq g(\lambda^{\text{fl}}) \), then \( \lambda^* = \lambda^{\text{fl}} \). Indeed, for all \( \lambda \), if \( g(\lambda) \) is active, then \( g(\lambda) \geq f(\lambda) \geq f(\lambda^{\text{fl}}) \).

Note that

\[
f(\lambda^{\text{fl}}) \geq g(\lambda^{\text{fl}}) \iff \mathbb{E}[(V - \lambda^{\text{fl}}^+)] - \psi \geq \max_{\phi} \{ (\phi - \lambda^{\text{fl}})\tilde{F}(\phi) \} \iff \lambda^{\text{fl}} \leq \lambda^{\text{nl}}.
\]

2. Similarly, if \( g(\lambda^{\text{nl}}) \geq f(\lambda^{\text{nl}}) \), then \( \lambda^* = \lambda^{\text{nl}} \). Using the fact that \( \lambda^{\text{nl}} = (\phi^*)^{-1}(\lambda^{\text{fl}}) \), we have

\[
g(\lambda^{\text{nl}}) \geq f(\lambda^{\text{nl}}) \iff \lambda^{\text{nl}} \geq \lambda^{\text{pl}} \iff \phi^*(\lambda^{\text{nl}}) \geq \phi^*(\lambda^{\text{pl}}) \iff \rho \leq \tilde{F}(\phi^*(\lambda^{\text{pl}})),
\]

where we have used the fact that \( \phi^*(\lambda) \) is an increasing function. Also, note that by the expression (EC.4), we have \( \phi^*(\lambda^{\text{pl}}) \geq \lambda^{\text{pl}} \).

3. In the last case, if \( f(\lambda^{\text{fl}}) > g(\lambda^{\text{fl}}) \) and \( g(\lambda^{\text{nl}}) < f(\lambda^{\text{nl}}) \), then \( \lambda^* = \lambda^{\text{pl}} \). This happens when \( \lambda^{\text{pl}} < \lambda^{\text{fl}} < \phi^*(\lambda^{\text{pl}}) \). □
EC.2.3. Proof of Proposition 4

We break up the proof into three cases, based on the value that $\lambda^*$ can take as per Lemma 5.

Case 1: $\lambda^* = \lambda^{f_{i}}$. We show that $\pi^{f_{i}}$ is optimal. As observed before, note that the choice of $\phi_{i}^*$ here is moot since the strategy learns on every $i$.

This strategy is feasible for (6), since by the definition of $\lambda^{f_{i}}$ inventory consumption under $\pi^{f_{i}}$ is

$$N \cdot P[V \geq \lambda^{f_{i}}] = N \cdot F(\lambda^{f_{i}}) = N \cdot \frac{K}{N} = K.$$  

Moreover, the objective function in (6) is equal to

$$N \cdot \left( \mathbb{E}[V \cdot 1\{V \geq \lambda^{f_{i}}\}] - \psi \right) = N \cdot \left( \mathbb{E}[V \cdot 1\{V \geq \lambda^{f_{i}}\}] - \psi \right) + K \cdot \lambda^{f_{i}} - N \cdot \lambda^{f_{i}} \cdot \frac{K}{N}$$

$$= N \cdot \left( \mathbb{E}[V \cdot 1\{V \geq \lambda^{f_{i}}\}] - \psi \right) + K \cdot \lambda^{f_{i}} - N \cdot \lambda^{f_{i}} \cdot \mathbb{P}[V \geq \lambda^{f_{i}}]$$

$$= N \cdot \left( \mathbb{E}[(V - \lambda^{f_{i}})^+] - \psi \right) + K \cdot \lambda^{f_{i}}$$

$$= J_{FB-MC}^{f_{i}}.$$  

Thus, since $\pi^{f_{i}}$ is feasible for both (6) and its relaxation (7) and $J_{FB-MC}^{f_{i}}$ upper bounds the firm first-best payoff, it must be an optimal solution to (6).

Case 2: $\lambda^* = \lambda^{n_{i}}$. In this case, the inventory constraint of (6) is also satisfied by definition of $\lambda^{n_{i}}$. Moreover, the solution yields the following objective function

$$N \cdot \phi^*(\lambda^{n_{i}}) \cdot F(\phi^*(\lambda^{n_{i}})) = N \cdot \phi^*(\lambda^{n_{i}}) \cdot F(\phi^*(\lambda^{n_{i}})) + K \cdot \lambda^{n_{i}} - N \cdot \lambda^{n_{i}} \cdot F(\lambda^{n_{i}})$$

$$= N \cdot (\phi^*(\lambda^{n_{i}}) - \lambda^{n_{i}}) \cdot F(\phi^*(\lambda^{n_{i}})) + K \cdot \lambda^{n_{i}}$$

$$= J_{FB-MC}^{n_{i}}.$$  

Since $J_{FB-MC}^{n_{i}}$ upper bounds $J_{FB-MC}^{f_{i}}$, this shows the optimality of $\pi^{n_{i}}$ by a similar argument to the one used in Case 1.

Case 3: $\lambda^* = \lambda^{p_{i}}$. Note that in this case, we have

$$F(\lambda^{p_{i}}) \leq 1 - \frac{K}{N} \leq F(\phi^*(\lambda^{p_{i}})) \implies F(\lambda^{p_{i}}) \geq \frac{K}{N} \geq F(\phi^*(\lambda^{p_{i}})).$$  

Consequently, there exists $\eta$ such that $\eta F(\lambda^{p_{i}}) + (1 - \eta) F(\phi^*(\lambda^{p_{i}})) = K/N$, which we set $q_{i}$ to. It is then straightforward that the firm solution satisfies the inventory constraint for (6) and yields an objective function equal to $J_{FB-MC}^{p_{i}}$, which shows $\pi^{p_{i}}$ is optimal for (6).

□
EC.2.4. Proof of Lemma 6

The proof again follows the template of the proof of Proposition 1. In particular, we can show that for $M$ large enough, (4) is feasible and can be formulated as follows

$$z^*(\lambda, M) = \min z$$

s.t. $z \geq \psi + \max_p \{t(p) \cdot \bar{F}(p)\}$

$$z = \mathbb{E}[t(V)]$$

$(EC.5)$

$$0 \leq t(p) \leq M, \quad \forall p$$

$$t(p) = 0, \quad \forall p \leq \lambda.$$ 

Since the proof is almost similar to that of Proposition 1, we only highlight the pair of primal/dual and their associated solution. Fixing $z$ to $z^*(M)$, the optimal contract can be found by solving the following problem:

$$\min y$$

s.t. $y \geq t(p) \cdot \bar{F}(p)$, $\forall p$

$$z^*(\lambda, M) = \mathbb{E}[t(V)]$$

(primal)

$$0 \leq t(p) \leq M, \quad \forall p$$

$$t(p) = 0, \quad \forall p \leq \lambda.$$ 

The dual problem can be written as

$$\max \zeta \cdot z^*(M) - M \cdot \int \xi(p) \cdot dp$$

s.t. $\int \nu(p) \cdot dp = 1$

$$\bar{F}(p) \cdot \nu(p) + \xi(p) \geq \zeta \cdot f(p), \forall p \geq \lambda$$

$$\nu(p) \geq 0, \quad \xi(p) \geq 0.$$ 

We now exhibit the pair of feasible primal/dual solutions for which the primal and dual objective functions are equal. The primal solution is defined as

$$t(p) = \begin{cases} 0, & \text{if } p \leq \lambda, \\
\min \left\{ M \cdot \frac{\bar{F}(\mu^*)}{F(p)}, M \right\}, & \text{otherwise.}
\end{cases}$$

where $\mu^*$ is chosen so that

$$\int_{p \geq \lambda} \min \left\{ M \cdot \frac{\bar{F}(\mu^*)}{F(p)}, M \right\} \cdot f(p) \cdot dp = z^*(\lambda, M).$$

The dual solution is defined as

$$\nu(p) = \begin{cases} f(p) \cdot C, & \text{if } \lambda \leq p \leq \mu^*, \\
\frac{f(p) \cdot C}{\bar{F}(p)}, & \text{if } p \leq \mu^*, \\
0, & \text{otherwise,}
\end{cases}$$

$$\xi(p) = \begin{cases} f(p) \cdot C, & \text{if } p \geq \mu^*, \\
0, & \text{otherwise,}
\end{cases}$$
where \( C \) is chosen such that
\[
\int_\lambda \mu^* \cdot d\nu(p) = 1.
\]

Proving that these two solutions lead to the same primal and dual objective functions proceed analogously to the proof of Proposition 1 and is thus omitted. We end this proof by highlighting an interesting property of the optimal contract \( t^M_{\lambda,\mu^*} \). Since the IC constraint must be binding, we must have
\[
z^*(\lambda, M) = \psi + \max_p \{ t^M_{\lambda,\mu^*} (p) \cdot \bar{F}(p) \},
\]
which in turns implies that
\[
E(t^M_{\lambda,\mu^*}(V)) - \psi = t^M_{\lambda,\mu^*}(p) \cdot \bar{F}(p) = \max_{\phi > 0} \{ t^M_{\lambda,\mu^*}(\phi) \cdot \bar{F}(\phi) \}, \quad \forall \lambda \leq p \leq \mu^*.
\]
\[
\text{(EC.6)}
\]

**EC.2.5. Proof of Lemma 7**

**Part 1.** We first show that \( (\pi^{f_i}, T^{f_i}) \) satisfies (IC-multi) and (participation-multi). Note that satisfying (IC-multi) automatically implies (participation-multi), since any uninformed price policy yields the agent a non-negative payoff.

We show that \( (\pi^{f_i}, T^{f_i}) \) satisfies (IC-multi), i.e. that \( \pi^{f_i} \in \arg \max \{ U(\pi, T^{f_i}) : \text{s.t. } \pi \text{ satisfies (inventory)} \} \). Recall that \( T^{f_i} = \{ t^{f_i} \} \) where \( t^{f_i} = t^M_{\lambda^{f_i},\mu^{f_i}} \), we now check that \( \pi^{f_i} \) satisfies (inventory). We have
\[
I(\pi^{f_i}) = \sum_i q_i^{f_i} \cdot E[1\{V_i \geq \theta_i^{f_i}(V_i)\}] + \sum_i (1 - q_i^{f_i}) \cdot \bar{F}(\phi_i^{f_i}) = N \cdot \bar{F}(\lambda^{f_i}) = K,
\]
where the second equality follows from the definition of of \( \pi^{f_i} \), and the third from the fact that \( \bar{F}(\lambda^{f_i}) = \rho = K/N \). Thus, \( \pi^{f_i} \) satisfies (inventory).

To conclude, we must also show that \( \pi^{f_i} \) maximizes \( U(\pi, T^{f_i}) \) among all inventory feasible policies. Recall that \( T^{f_i} = \{ t^{f_i} \} \) where \( t^{f_i} = t^M_{\lambda^{f_i},\mu^{f_i}} \). For any arbitrary inventory feasible policy \( \pi = (q_i, \theta_i, \phi_i)_{i=1}^N \in \Pi(K) \),
\[
U(\pi^{f_i}, T^{f_i}) = \sum_i q_i^{f_i} \cdot (E[t^{f_i}(\theta_i^{f_i}(V_i)) \cdot 1\{V_i \geq \theta_i^{f_i}(V_i)\}] - \psi) + \sum_i (1 - q_i^{f_i}) \cdot t^{f_i}(\phi_i^{f_i}) \cdot \bar{F}(\phi_i^{f_i})
\]
\[
= N \cdot (E[t^{f_i}(V)] - \psi)
\]
\[
= \sum_i q_i \cdot (E[t^{f_i}(V)] - \psi) + \sum_i (1 - q_i) \cdot \max_{\phi_i} \{ t^{f_i}(\phi_i) \cdot \bar{F}(\phi_i) \}
\]
\[
\geq \sum_i q_i \cdot (E[t^{f_i}(\theta_i(V))] 1\{V_i \geq \theta_i(V_i)\}] - \psi) + \sum_i (1 - q_i) \cdot t^{f_i}(\phi_i) \bar{F}(\phi_i)
\]
\[
= U(\pi, T^{f_i}),
\]
where in the third equality we have used the fact that by (EC.6) in the proof of Lemma 6, \( \mathbb{E}[T^{i,\cdot}] - \psi = \max_{\phi_i} \{ t^{i,\cdot}(\phi_i) \cdot F(\phi_i) \} \), and in the first inequality, the facts that \( t^{i,\cdot}(V_i) \geq t^{i,\cdot}(\theta_i(V_i)) \cdot 1 \{ V_i \geq \theta_i(V_i) \} \) and that \( t^{i,\cdot}(\cdot) \) is non-decreasing. Since \( \pi^{t^{i,\cdot}} \) then achieves at least as high an objective as any policy in \( \Pi(K) \) and is feasible, it must be optimal for the agent’s control problem and thus incentive compatible.

**Part 2.** We now turn to the other case, when the firm wants to induce partial learning, and show that \( (\pi^{p^i}, T^{p^i}) \) satisfies \( \text{(IC-multi)} \) and \( \text{(participation-multi)} \). The proof proceeds in a similar manner to that of the previous case. We first ensure this policy is inventory feasible:

\[
I(\pi^{p^i}) = \sum_i q_i^{p^i} \cdot \mathbb{E}[1\{ V_i \geq \theta_i^{p^i}(V_i) \}] + \sum_i (1 - q_i^{p^i}) \cdot \tilde{F}(\phi_i^{p^i}) \\
= \sum_i \left( \eta_i \cdot \mathbb{E}[1\{ V_i \geq \lambda_i^{p^i} \}] + (1 - \eta_i) \cdot \tilde{F}(\phi^*(\lambda_i^{p^i})) \right) \\
= K,
\]

where in the last line we used the definition of \( \eta \) from Proposition 4.

We then show that \( \pi^{p^i} \) maximizes \( U(\pi, T^{p^i}) \) among all feasible policies. We do this by proving that it achieves at least the objective of any other feasible policy \( \pi = (q_i, \theta_i, \phi_i)_{i=1}^N \). Recall that \( T^{p^i} = \{ t^{p^i} \} \) where \( t^{p^i} = t_{\lambda_i, t, \mu_i}^{p^i} \). We have:

\[
U(\pi^{p^i}, T^{p^i}) \\
= \sum_i q_i^{p^i} \cdot (\mathbb{E}[t^{p^i}(\theta_i^{p^i}(V_i)) \cdot 1\{ V_i \geq \theta_i^{p^i}(V_i) \}] - \psi) \\
+ \sum_i (1 - q_i^{p^i}) \cdot t^{i,\cdot}(\phi_i^{p^i}) \cdot \tilde{F}(\phi_i^{p^i}) \\
= N \cdot \eta \cdot (\mathbb{E}[t^{p^i}(V_i) \cdot 1\{ V_i \geq \lambda_i^{p^i} \}] - \psi) + N \cdot (1 - \eta) \cdot t^{p^i}(\phi^*(\lambda_i^{p^i}) \cdot \tilde{F}(\phi^*(\lambda_i^{p^i})) \\
= N \cdot \eta \cdot (\mathbb{E}[t^{p^i}(V_i)] - \psi) + N \cdot (1 - \eta) \cdot t^{p^i}(\phi^*(\lambda_i^{p^i})) \cdot \tilde{F}(\phi^*(\lambda_i^{p^i})) \\
\geq \sum_i q_i \cdot (\mathbb{E}[t^{p^i}(\theta_i(V_i)) \cdot 1\{ V_i \geq \theta_i(V_i) \}] - \psi) + \sum_i (1 - q_i) \cdot t^{i,\cdot}(\phi_i) \cdot \tilde{F}(\phi_i) \\
= U(\pi, T^{p^i}).
\]

In the above, the third equality follows by definition of \( t^{p^i} \). Next, observe that by definition of \( \phi^*(\lambda_i^{p^i}) \), \( \lambda_i^{p^i} \leq \phi^*(\lambda_i^{p^i}) \leq \mu_i \), and thus using by (EC.6) in the proof of Lemma 6 we have that \( \mathbb{E}[T^{p^i}] - \psi = t^{p^i}(\phi^*(\lambda_i^{p^i})) \cdot \tilde{F}(\phi^*(\lambda_i^{p^i})) \). This guarantees that the fourth equality holds. Furthermore, since (EC.6) also implies that \( t^{p^i}(\phi^*(\lambda_i^{p^i}) \cdot \tilde{F}(\phi^*(\lambda_i^{p^i})) = \max_{\phi_i} \{ t^{p^i}(\phi_i) \cdot \tilde{F}(\phi_i) \} \geq t^{p^i}(\phi_i) \cdot \tilde{F}(\phi_i) \).

Since \( t^{p^i}(\cdot) \) is non-decreasing, \( t^{p^i}(V_i) \geq t^{p^i}(\theta_i(V_i)) \) for \( \theta_i(V_i) \leq V_i \) and thus the first inequality follows. As observed in Part 1, this also implies \( \text{(participation-multi)} \).
EC.2.6. Proof of Lemma 8

The proof of this lemma proceeds analogously to that of Proposition 2 and is thus omitted.

EC.2.7. Proof of Corollary 5

Let us first assume that \( J^{FB-MC} = R(\pi^{FB}) - \psi \) is achieved by \( \pi^{FB} \in \{\pi^{f,l}, \pi^{p,l}\} \). By Lemma 8, for any \( \epsilon > 0 \), there exists \( M_{\epsilon} \) and a contract \( T_{\epsilon} \) such that

\[
P(\pi^{FB}, T_{\epsilon}) \leq \psi + \epsilon. \tag{EC.7}
\]

Moreover, by Lemma 7, the pair \( (\pi^{FB}, T_{\epsilon}) \) satisfies \((\text{IC-multi})\) and \((\text{participation-multi})\). Thus the objective of the principal is:

\[
R(\pi^{FB}) - P(\pi^{FB}, T_{\epsilon}) \geq R(\pi^{FB}) - \psi + \epsilon = J^{FB-MC} + \epsilon,
\]

where in the inequality we have used equation (EC.7).

In the case that \( J^{FB-MC} \) is achieved by \( \pi^{FB} = \pi^{n,l} \), then the principal can just set \( T_{\epsilon} = \{\epsilon \cdot 1\{p = \phi^*(\lambda^{\text{-informed}})\}\} \). In that case,

\[
R(\pi^{FB}) - P(\pi^{FB}, T_{\epsilon}) = R(\pi^{FB}) - \epsilon.
\]

\( \square \)

EC.3. Likelihood ratio does not satisfy MLRP

We discuss how the MLRP condition typically used in the analysis of principal-agent problems fails in our case. In particular, we discuss two natural counterparts of the MLRP condition for our setting.

In the binary effort case, if \( h(\cdot|e) \) were the density of the revenue \( Y \) that the firm garners when the learning effort is \( e \), then a natural counterpart of the classical MLRP condition would be that

\[
[h(y|e = 1) - h(y|e = 0)]/h(y|e = 1)
\]

is increasing. This is ill-defined in our setting since \( h(y|e = 0) \) depends on the best uninformed price which in turns depends on the contract.

Another natural counterpart of the classical MLRP, would be to assume that there is a continuous space of effort level wherein we would create an effort level for each uninformed price when there is no learning. In this case, we could check the MLRP condition by inspecting the ratio \([h(y|e = 1) - h(y|e = 0, \phi)]/h(y|e = 1)\) where \( h(\cdot|e = 0, \phi) \) denotes the density of the revenue \( Y \) that the firm garners when the learning effort is \( e = 0 \) and the uninformed price is \( \phi \). Then

\[
\frac{h(y|e = 1) - h(y|e = 0, \phi)}{h(y|e = 1)} = \frac{f(y) - \tilde{F}(y)1\{y = \phi\}}{f(y)} = 1 - \frac{\tilde{F}(y)1\{y = \phi\}}{f(y)}.
\]

Clearly, this ratio is not monotonic in \( y \), and not strictly monotonic over subsets of the range of \( y \), which is assumed via the MLRP condition.