Learning Deterrence vs. Encouragement: Seller-optimal Buyer-learning

Qianjun Lyu*

[Preliminary]

March 1, 2021

Abstract

This paper studies the optimal selling mechanism when an uninformed buyer can sequentially learn about his value of a product privately. The seller designs the mechanism to affect the buyer's benefit from learning and thereby controls the learning process. Our main result shows that it is always sub-optimal to induce partial learning, and the optimal mechanism either encourages full learning or deters the buyer from private learning. The optimality between learning deterrence and encouragement depends on the buyer's prior belief, which is a measure of both the initial informativeness and the level of optimism. Especially, the seller optimally encourages full learning if the buyer is relatively uninformative but optimistic. Otherwise, learning is deterred and the buyer is strictly better off if the mechanism deters learning. When the cost of learning converges to zero, our results tribute to the well-known niche-mass market analysis.

Key words: buyer-learning, sequential information acquisition, return policy, learning deterrence, information design.

*Faculty of Business and Economics, University of Hong Kong, lyu2014@hku.hk.

1 Introduction

Consumers gathering information about a product is universal nowadays. Apart from visiting the off-line stores to physically experience the product, they can also search for customer reviews and professional evaluations shared by individual users on the Internet. Hence, the seller may not be able to fully control the buyer's information sources. However, by designing a selling mechanism to affect the buyer's benefit from learning, the seller can indirectly control the buyer's endogenous learning. In this paper, we study the seller's revenue-maximizing selling mechanism anticipating that the buyer privately performs sequential learning about his value of a product.

A selling mechanism specifies the selling price and the return policy.[1] In general, return policies vary across platforms and product categories. Free return (full refund) and no return (non-refundable payment) are the most common return policies. Apart from these, airline companies usually charge a fixed fee for ticket refund and the fees vary across different travel class. Fashion platform like Farfetch offers free return and free home collection, while others, like Luisavaroma, offer free return only in the form of store credit. More surprisingly, Amazon offers refund with no return, however it is random at the moment of purchasing and the company specifies it as "Amazon" *may* determine that a refund can be issued without requiring a return". Despite the various formats of return policies, we can represent a return policy as the probability that a return is required (by the seller) and the amount of refund,[2] since the consumer's preference is quasi-linear.

To be concrete, we consider a seller (she) selling one unit of indivisible good to a single buyer (he), who is initially uninformed about his true value of the product, which could be either high or low with $v_h > v_l > 0$. The buyer's prior belief is commonly known, but he can update his belief by performing exponential experiment sequentially. In particular, by exerting costly effort, e.g., spending time acquiring information, a good news arrives with rate λ if his true value is high, while no news arrives if his true value is low.[3] The seller obtains zero utility from keeping the good and there is no production cost or return cost. At the beginning of the time, the seller commits to a time-persistent selling mechanism.[4] The buyer then decides how much information to

[1] This can be verified to be without loss of generality.

[2] The refund is issued independent of whether a return is required.

[3] The reader can interpret no news as bad news, e.g., high value can generate both good signal and bad signal, while low value can only generate bad signal.

[4] In practical, the retailer usually imposes a limit window of time when a return can be accepted—

acquire and makes the purchase/return decision based on the information outcomes.

Our model builds on the exponential-bandit framework (Keller, Rady and Cripps (2005)), which grants us significant tractability and allows us to frame this mechanism design problem as an information design problem. Loosely speaking, we can characterize the (incentive compatible) selling mechanism as a function of buyer's posterior belief and ask what is the seller-optimal amount of buyer-learning. Basically, the seller maximizes her expected revenue over the distributions of buyer's stopping beliefs.

To do this, we claim that for a fixed selling price, the buyer's expected trading surplus given any optimal mechanism is the same with his continuation value under a no return mechanism. The latter one is a standard stopping time problem,[5] and the net surplus of high value buyer—v_h minus price—is a sufficient statistics in characterizing the buyer's value function. Therefore, while picking up the optimal selling mechanism, the seller first chooses the selling price to affect the level of information value that the buyer can attain through learning, and then design a return policy to truncate the buyer's sequential learning. Essentially, the return policy compensates the buyer with the opportunistic information rent—continuation value—that he can enjoy if he continues to learn,[6] and thereby enforces earlier stopping. For a given price, incorporating buyer's optimality (also known as smooth-pasting) of stopping learning and requesting a return, we can pin down the return policy as a function of the buyer's stopping belief.

Before we elaborate the results, it would be useful to introduce three types of selling mechanism: (1) Learning Deterrence; (2) Stochastic Return; and (3) Free Return.

Learning Deterrence refers to a particular *no return* mechanism such that the selling price is designed to make the buyer indifferent between purchasing the item immediately and performing learning. In other words, the buyer obtains trading surplus equal to the opportunistic information rent that he can attain through learning. Given the

same return policy persists within the return window and no return is allowed afterwards. In this paper, we omit return window for simplicity. But given our information acquisition process, the buyer always makes decision in a finite period of time, hence this simplification will not affect the results crucially since the retailer can choose the return window to make sure the current optimal selling mechanism works. But our results may not speak to the situations where the seller can commit to a time-contingent return protocol.

[5]Essentially, given a no return mechanism, the buyer performs all learning before trading. If a good news arrives, his belief jumps to 1 and he purchases it. If no news arrives for a sufficient amount of time, he walks away.

[6]Whenever the buyer feels indifferent, we let him break the tie in terms of the seller's favor.

seller-preferred tie-breaking rule, learning is deterred and allocation is efficient. It turns out that the selling price of Learning Deterrence is the lowest among these three mechanisms, which indicates that the buyer is strictly better off if the mechanism deters learning rather than encourages learning.

Stochastic Return stands for the type of selling mechanisms with a random return policy—when a buyer requests a return, he can keep the item with positive probability but receives some refund for sure. Hence, the seller is able to induce some particular level of *partial learning* in the sense that the buyer optimally stops learning and requests a return even though the continuation value for continuing to learn is still *positive*. Meanwhile, trading surplus is positive upon return, which ensures the seller a positive return revenue.

Free Return allows the buyer to opt out freely, hence the seller can encourage *full learning* in the sense that the buyer requests a return when the continuation value from keeping learning becomes *zero*. Moreover, trading is ex-ante inefficient and the seller obtains zero sale revenue.

In general, the above three mechanisms are sufficient to induce all feasible learning-dynamics—buyer's posterior distributions—under our assumption of exponential experiment. Meanwhile, Learning Deterrence and Free Return can be interpreted as the opposite limit of Stochastic Return in terms of how much information the buyer learns. After the seller picks the price, anyone of these three mechanisms could be optimal.[7] However, our main result shows that Stochastic Return is always sub-optimal. To be more concrete, for a given Stochastic Return mechanism, the seller is strictly better off by *either* lowering the price and preventing the buyer from private learning *or* raising the price and encouraging full learning. That is, the revenue-maximizing mechanism is either Learning Deterrence or Free Return.

Given this, we fully characterize the optimal selling mechanism, which depends the buyer's prior belief. The buyer's prior belief is a measure of both his initial informativeness and the level of optimism. Being less informed ex-ante indicates greater information value that the buyer can attain through learning. Therefore, to deter learning, the seller has to compensate the buyer's opportunistic information value, which is higher when the buyer is initially more uncertain. Conversely, with Free Return,

[7] Precisely, there is one optimal return policy associate with each selling price. When the price varies, the corresponding optimal return policy could induce no learning, partial learning, or full learning.

the seller encourages the buyer to learn, which avoids her to pay for the opportunistic information value. However, Free Return causes allocation inefficiency, which is less severe when the buyer is more optimistic initially. Essentially, with a higher prior belief, there is larger chance of realizing a good news and the seller can induce a more decent chance of successful sale. Meanwhile, more optimistic prior indicates higher willingness to pay, which also results in a relatively higher price. Above all, the seller optimally picks Free Return when the buyer is initially more uncertain but also more optimistic. Otherwise, Learning Deterrence is optimal. It might be counter-intuitive since the seller deters learning for both extremely optimistic and pessimistic buyer. This is because for those buyers, they are already very informed initially and the opportunistic information value is almost zero. Hence, the buyer is willing to give up his option to learn even with very small trading surplus, then the seller can simply attain almost full allocation surplus.

With the optimal selling mechanism, the selling price and the seller's expected revenue are increasing in the buyer's prior belief. However, the buyer's ex-ante payoff is not monotone in his prior. Considering merely Learning Deterrence, the buyer's expected payoff is just his continuation value from learning, which is a bell-shaped curve on his prior belief. Regarding the set of prior within which the seller optimally chooses Free Return, the joint trading surplus is smaller compared to Learning Deterrence since allocation is not efficient and there is a sunk cost for learning. Meanwhile, the seller's expected revenue is larger than with Learning Deterrence, which leaves the buyer with a strict reduction in his own trading surplus.

Furthermore, as learning becomes more efficient, e.g., the cost of learning is smaller, the set of prior belief that the seller optimally selects Free Return expands. When the cost of learning converges to zero, the buyer can almost learn perfect information. Therefore, with Free Return, the seller sets the price arbitrarily close to v_h and let go the buyer almost sure to have low value. With Learning Deterrence, she sets price arbitrarily close to v_l intending to cover the whole market. This tributes to the standard analysis of niche-mass market. The ratio v_l/v_h pins down the cutoff of prior where the seller feels indifferent.

Related literature. There is a growing literature on mechanism design incorporating information as part of optimal choice. Considering price discrimination, Li and Shi (2017) allow the seller to disclose different additional information for different types of buyer. They show that partial and discriminatory disclosure weakly dominates full disclosure in terms of seller's revenue. Guo, Li and Shi (2020) then characterize the property of optimal discriminatory disclosure. Bergemann and Pesendorfer (2007) allow the buyer to acquire information but the information accuracy is controlled by the seller. Johnson and Myatt (2006) introduce rotations of demand curves to capture the dispersion of consumer valuations and discuss how the seller's profit changes with the levels of dispersion. In stead of letting the seller optimally make restrictions on the buyer's learning process, Roesler and Szentes (2017) allow the buyer to acquire fully flexible and cost-less information, anticipating that the seller's pricing decision depends on his expected information outcomes. Ravid, Roesler and Szentes (2019) discuss the same scenario but the seller and buyer move simultaneously.

Our paper stays in the line of mechanism design when the buyer can endogenously acquire information (Shi (2012) and Mensch (2020)), meaning that the seller uses different mechanisms to control the buyer's optimal information acquisition. Shi (2012) adopts rotational-ordered information technology and shows the optimality of posted price in the case of single buyer. Moreover, the optimal price, conditional on a fixed information choice being the equilibrium outcome, is smaller than the monopoly price when this information choice is exogenous. This is because the price increment affects the buyer's incentive to acquire information and hence changes the trading probability, which turns out to lower the seller's expected revenue. However, in our paper, by assuming dynamic information acquisition, it is actually the return policy that pins down the buyer's optimal stopping and the seller can sustain the same trading probability when price increases.

Mensch (2020) discusses the same question but allows flexible information acquisition with cost as the expected difference in a posterior-separable measure of uncertainty. It characterizes the set of all implementable mechanisms (*contour mechanisms*), which consist of triplets of allocation probabilities, prices and beliefs. Therefore, the problem can be translated into Bayesian persuasion. Such translation is also adopted in our model. However, we choose instead exponential bandit as information technology and assume additive time cost, therefore the cost for the same Blackwell experiment will be the same for different prior beliefs, which is not true for flexible information.[8] This allows us to discuss how the buyer's prior belief (ex-ante informativeness) affects the seller's optimal selling mechanism and in turn his own trading surplus.

[8]There does not exist a unified measure of uncertainty, regardless of the prior belief, that can represent the additive time cost of Poisson signal (See Appendix A of Mensch (2020) and Pomatto, Strack and Tamuz (2019)).

Many other papers also discuss seller's pricing strategy when the buyer can acquire information dynamically (Bonatti (2011), Bergemann and Valimaki (2000) and Bergemann and Valimaki (1996)). The most related papers are Lang (2019) and Pease (2018). Both consider the seller's optimal pricing when the buyer can sequentially acquire information before purchasing. Therefore, the buyer's optimal stopping only responds to the selling price. However, in our paper, by introducing the return policy, the seller can induce more flexible stopping time and obtain non-negative transfer upon the buyer's stopping. In Lang (2019), buyer's information acquisition follows Brownian motion whose drift depends on the buyer's true valuation, which is normally distributed. He shows that the trading probability is increasing in the buyer's prior valuation (mean). However, we show that allocation efficiency is non-monotone in buyer's prior belief. The major difference is that, in our model, the buyer's prior belief (probability of having high valuation) captures his initial informativeness. Hence as long as the buyer is sufficiently optimistic or pessimistic, the potential gain from acquire information becomes small, which give rise to efficient trading. While in Lang (2019), the buyer's prior mean represents his ex-ante valuation of the product, which does not affect the informativeness of the buyer's prior belief.

This paper is also related to Armstrong and Zhou (2015). They study price discrimination where the buyer can subsequently search for outside options. This motivates the seller to offer a buy-now discount at the first meeting with the buyer to deter search and thereby achieve immediate sale. We apply the similar logic when discussing the optimality of Learning Deterrence, nevertheless our buyer is uninformed about the product itself rather than his outside options. Therefore, when we vary the buyer's initial informativeness, Learning Deterrence could be too expensive to implement and the seller optimally selects Free Return. Our result that the seller's optimal selling mechanism either induces full learning or no learning is close to Johnson and Myatt (2006), which shows that the seller's profit is quasi-convex in the level of demand dispersion. However, their seller can directly choose the distribution of consumer valuations, which, in our paper, is the consumer's best response to the selling mechanism.

The remainder of this paper is organized as follows: Section 2 discusses the model and Learning Deterrence. Section 3 studies Learning encouragement. Section 4 characterizes the optimal selling mechanism. Section 5 gives a discussion. Section 6 concludes.

2 Model

A seller sells one unit of indivisible good to a risk-neutral buyer. The product values zero to the seller and there is no cost of production and return. The buyer is initially uninformed about his true value of the product, which is either high or low with $v_h > v_l > 0$. Let μ be the buyer's belief of having high value and we call this as the buyer's type. Type-μ buyer's expected value is $\mathbb{E}(v|\mu) = \mu v_h + (1-\mu)v_l$. The buyer's initial type μ_0 is common knowledge and his posterior belief evolves over time conditional on his learning dynamic, which will be specified later. We use τ to denote the time and $\mu(\tau = 0) = \mu_0$. The seller commits to a selling mechanism, which specifies (1) a selling price $t_b > 0$ which is the transfer made from buyer to the seller at the time of purchasing;[9] and (2) a return policy that describes: a) the probability that the buyer is required to return the item at the time of requesting a return, and b) the refund that the buyer obtains regardless of whether a return is required.[10] Equivalently, we can use (x_r, t_r) to denote the return policy.[11] $x_r \in [0, 1]$ is the probability that a buyer keeps the item after requesting a return and $t_r \in [0, t_b]$ is the net transfer made from buyer to seller at the time of requesting a return.[12] Overall, one typical selling mechanism is characterized by $\{t_b, (x_r, t_r)\}$. Given this, a "No Return" mechanism can be written down as $\{t_b, (1, t_b)\}$. "Free Return" can be represented as $\{t_b, (0, 0)\}$. We define Stochastic Return as $\{t_b, (x_r, t_r)\}$ with $x_r \in (0, 1)$. It is without loss to assume $v_h - t_b > v_h x_r - t_r$. That is, the type-1 buyer purchases the item without requesting further return. Denote $s = v_h - t_b$ as the net surplus obtained by high value buyer. For simplicity, we assume both parties have no discount over time. The buyer's outside option is normalized to be zero.

A type-μ buyer's payoff is realized when he consumes the item.[13] And if so, he cannot

[9] The seller can offer a pair of (x_b, t_b), which specifies the probability of sale x_b and the transfer t_b at the time of purchasing. We show that it is without loss to focus on $x_b = 1$.

[10] It refers to the Amazon example of "refund without requiring a return". Besides, in some cases, the refund could depend on whether the buyer returns the item. Due to quasi-linear payoff structure, it is without loss to focus on the expected refund.

[11] In general, the seller can design a vector of return policies to screen different types of buyers. It is without loss to focus on one piece of return policy.

[12] To understand it, if the buyer requests a return under the return policy (x_r, t_r), then the seller issues a refund $t_b - t_r$ to the buyer and at the meantime she rolls a dice. With probability $1 - x_r$, the buyer has to return the item to the seller, while with probability x_r, he can still keep it.

[13] The reader can consider that consuming the item provides free and perfect information about the true value.

request a return regardless of the return policy. In particular, type-μ buyers obtains utility $\mathbb{E}(v|\mu) - t_b$ if he purchases a non-refundable item, while he obtains utility $\mathbb{E}(v|\mu)x_r - t_r$ if he requests a return at policy (x_r, t_r). Let $\mathbf{B}_\tau$ be the indicator function of whether a purchase is occurred up to and *including* time τ. Hence, the time of purchasing is $\tau_b = \min\{\tau : \mathbf{B}_\tau = 1\}$. In an analogical manner, $\mathbf{R}_\tau$ denotes the indicator function of whether a return is occurred till time τ and the time of requesting a return is $\tau_r = \min\{\tau : \mathbf{R}_\tau = 1\}$. Naturally, $\tau_r \geq \tau_b$. The seller's revenue is denoted as Π.

$$\Pi = \mathbb{E}\left[\int_0^\infty t_b d\mathbf{B}_\tau + (t_r - t_b) d\mathbf{R}_\tau\right] \tag{1}$$

In this paper, we adopt the exponential-bandit framework. In particular, if the buyer decides to learn, he needs to pay a flow cost k. A good news arrives at Poisson rate λ if his true value is v_h and no news arrives if his true value if v_l. For simplicity, we assume the information process is the same both before he purchases the item or after. When a good news arrives, the buyer's belief jumps to 1. While if no Poisson jump arrives, the buyer's belief evolves according to the following law of motion:

$$\mu'(\tau) = -\mu(\tau)(1 - \mu(\tau))\lambda < 0$$

Under a "No Return" mechanism $\{t_b, (1, t_b)\}$, the buyer performs all learning before purchasing the item and his optimal information acquisition in this scenario serves as a building block in developing the main results. We denote $V^0(\mu(\tau), s)$ as the buyer's value function when he faces a "No Return" mechanism with the selling price equal $v_h - s$. It is characterized by the Bellman equation below.

$$\begin{aligned} V^0(\mu(\tau), s) = \max\{\ 0\ ,\ &\mathbb{E}(v|\mu(\tau)) - (v_h - s), \\ &- kd\tau + \mu(\tau)\lambda d\tau s + (1 - \mu(\tau)\lambda d\tau)V^0(\mu(\tau + d\tau), s)\} \end{aligned} \tag{2}$$

At time τ, the buyer walks away with payoff 0, or purchases the item to get the expected consumption utility $\mathbb{E}(v|\mu(\tau)) - (v_h - s)$. If the buyer continues to learn for a instant of time $d\tau$, with probability $\mu(\tau)\lambda d\tau$, a Poisson jump arrives and he purchases the item obtaining net surplus s; with the remaining probability, no news arrives and his belief reduces to $\mu(\tau + d\tau)$ and the corresponding value is $V^0(\mu(\tau + d\tau), s)$. Conditional on learning, the Bellman equation leads to the differential equation below.

$$(1 - \mu)\mu\lambda V'(\mu, s) + \mu\lambda V(\mu, s) = \mu\lambda s - k \tag{ODE}$$

Recall from Keller, Rady and Cripps (2005), for a fixed s, there are two cutoff beliefs, the quitting belief $q(s)$ and the trial belief $Q(s)$ with $q(s) \leq Q(s)$,[14] that pin down the buyer's optimal learning strategy: he continues to learn when his belief falls in between, otherwise, he does not learning. Essentially,

$$q(s) = \{\mu : V_1(\mu, s) = 0 \text{ and } V(\mu, s) = 0\} \tag{3}$$

$$Q(s) = \{\mu : V(\mu, s) = \mathbb{E}(v|\mu) - (v_h - s)\} \tag{4}$$

where $V_1(\mu, s)$ denotes the partial derivative w.r.t the first argument. From (3), $q(s) = \frac{k}{\lambda s}$ and we can obtain the buyer's continuation value for learning:[15]

$$V(\mu, s) = \frac{1}{\lambda}[-k + \lambda\mu s - k(1-\mu)\log(\frac{\mu/(1-\mu)}{k/(\lambda s - k)})] \tag{5}$$

This construction involves one implicit assumption: when the buyer stops learning at $q(s)$, he indeed prefers to quit the market rather than accepting the price.

$$\mathbb{E}(v|q(s)) - (v_h - s) \leq 0 \tag{Learning-Feasibility}$$

We call this the *Learning-Feasibility* constraint. Since if it fails, regardless of the buyer's prior belief, no learning can be induced at all. To avoid trivial result, we assume there exists two distinct roots $\underline{s} < \overline{s}$ that this constraint binds.

Proposition 1. (Optimal buyer learning under $\{v_h - s, (1, v_h - s)\}$)
(1) If $s < \underline{s}$ or $s > \overline{s}$, the buyer optimally chooses not to learn and

$$V^0(\mu, s) = \max\{0, \mathbb{E}(v|\mu) - (v_h - s)\}$$

(2) When $s \in [\underline{s}, \overline{s}]$, if $\mu_0 \in [q(s), Q(s))$, the buyer exerts learning and optimally stops till either a signal arrives or belief falls below $q(s)$.

$$V^0(\mu, s) = \begin{cases} 0, & \mu < q(s) \\ V(\mu, s), & q(s) \leq \mu < Q(s) \\ \mathbb{E}(v|\mu) - (v_h - s), & \mu \geq Q(s) \end{cases}$$

If $\mu_0 < q(s)$ or $\mu_0 \geq Q(s)$, the buyer does not learn and

$$V^0(\mu, s) = \max\{0, \mathbb{E}(v|\mu) - (v_h - s)\}$$

[14] Since the buyer's belief goes down if no news arrives, therefore we call the higher cutoff belief as the trial belief.

[15] The general solution of the (ODE) is $V(\mu, s) = s - \frac{k}{\lambda}[1 + (1-\mu)\log(\frac{\mu}{1-\mu})] + (1-\mu)c$. Applying this into (3), we can get $c = -s + \frac{k}{\lambda}\log(\frac{k}{\lambda s - k})$ and $q = \frac{k}{\lambda s}$.

Figure 1 is a graphical illustration of Proposition 1. As mentioned above, for sufficiently high $s > \bar{s}$ and sufficiently low $s < \underline{s}$, the buyer optimally chooses not to learn regardless of his prior. The upper and lower gray curves in Figure 1 capture the buyer's value function in these two scenarios, which indicates that type-μ buyer optimally chooses between buying the good and walking away. With moderate $s \in [\underline{s}, \bar{s}]$, the standard results of exponential bandit apply. For example, see the red curve in Figure 1. When the buyer's prior belief falls into $[q(s), Q(s))$, he optimally performs learning till either a good news arrives and he purchases the item or no news arrives and he quits the market at $q(s)$.

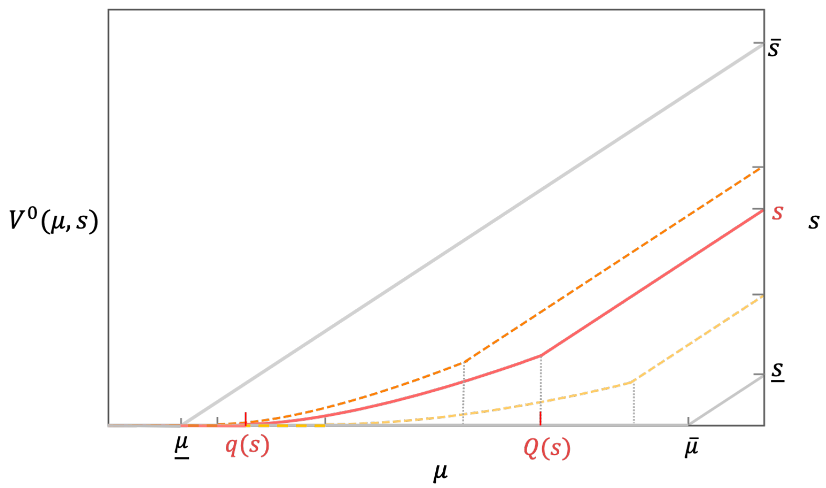

Figure 1: Optimal learning dynamic under "No Return" mechanisms

Note that s is a sufficient statistics in determining the level of the buyer's value function. When the seller raises s, the buyer's value function shifts up. Moreover, the learning interval $[q(s), Q(s))$ changes with s. It is easy to see that the quitting belief $q(s)$ is decreasing in s since the buyer optimally performs longer learning if a Poisson jump induces larger surplus. Meanwhile, the trial belief $Q(s)$ is also decreasing in s. Recall that $Q(s)$ is the cutoff belief that the buyer is indifferent between accepting the price and performing learning. When s increases by one unit, the price then decreases by one unit, hence the buyer's expected payoff from purchasing increases by one unit. However, if the buyer chooses to learn, the probability of obtaining a Poison jump is strictly less than 1 and thereby the increment of continuation value is less than one unit. Hence, in Figure 1, $[q(s), Q(s))$ shifts to the left while increasing s. More interestingly, the quitting belief and trial belief coincide at $\bar{s}$ and $\underline{s}$. Besides, the difference between the trial belief and quitting belief is single-peaked in s. Denote $\underline{\mu} = q(\bar{s})$ and $\bar{\mu} = q(\underline{s})$. In Figure 2, we plot the quitting belief and trial belief against s.

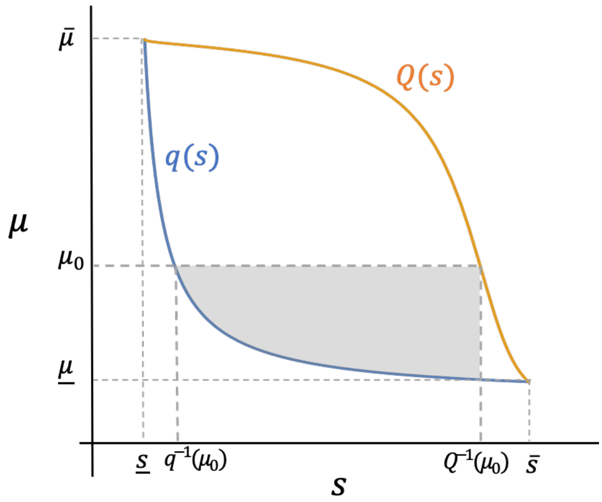

Figure 2: Feasible buyer-learning

Lemma 1. *For a fixed selling price $v_h - s$, all optimal selling mechanisms provide the buyer with ex-ante trading surplus equal $V^0(\mu_0, s)$.*

The formal proof is shown in the Appendix. To see the intuition, suppose the seller intends to set a predatory return policy (small x_r, large t_r) such that the buyer would strictly prefer to perform all learning before trading and only purchase when he obtains a good news, then such return policy cannot be realized. Lemma 1 holds trivially. Conversely, if the seller instead offers a benevolent return policy (large x_r, small t_r) intending to reward the buyer from early purchase, i.e., provides type-μ_0 buyer strictly higher payoff than $V^0(\mu_0, s)$. Then the seller can attain larger expected revenue by taking away part of this reward. Essentially, she can increase the buyer's chance of keeping the item upon return and charge a higher return transfer in a way such that the buyer preserves the same optimal stopping.

Corollary 1. *If $\mu_0 \notin [\underline{\mu}, \overline{\mu}]$, the optimal selling mechanism is $\{\mathbb{E}(v|\mu_0), (1, \mathbb{E}(v|\mu_0))\}$. The trading happens with probability 1 and the seller captures all trading surplus.*

This corollary describes an extreme case where the buyer is initially informative enough and he deems learning as sub-optimal. Hence, the seller can simply extract all trading surplus. When the buyer has a more uncertain prior belief, $\mu_0 \in [\underline{\mu}, \overline{\mu}]$, learning becomes valuable. If the seller intends to deter buyer's learning, she has to give away part of the trading surplus to compensate the buyer's *opportunistic information rent*, which we refer by the buyer's loss from not acquiring information. Consider Figure

2,[16] by setting $s \geq Q^{-1}(\mu_0)$, type-μ_0 buyer weakly prefers to purchase the item rather than acquiring information. Therefore, in terms of deterring learning, the seller lets the inequality bind to maximize revenue. Denote $t^D(\mu_0) := v_h - Q^{-1}(\mu_0)$. We call the mechanism $\{t^D, (1, t^D)\}$ as *Learning Deterrence*, with which allocation is efficient and learning is deterred.[17]

Proposition 2. (Learning Deterrence) *If* $\mu_0 \in [\underline{\mu}, \overline{\mu}]$ *and conditional on* $s \geq Q^{-1}(\mu_0)$, $\{t^D, (1, t^D)\}$ *is optimal. With which, the joint surplus is* $\mathbb{E}(v|\mu_0)$.
(1) The buyer obtains (weakly) positive trading surplus $V(\mu_0, Q^{-1}(\mu_0))$. *It is single-peaked in* μ_0 *and equals zero at the two end points;*
(2) The seller's revenue $t^D(\mu_0)$ *is increasing in* μ_0 *and equals* $\mathbb{E}(v|\mu_0)$ *at the two end points.*

To deter buyer's private learning, the seller has to sufficiently lower the price so that accepting the price is more attractive for the buyer compared to learning. $t^D(\mu_0)$ is the highest achievable price to prevent type-μ_0 buyer from learning. Figure 3 depicts the allocation of joint surplus under Learning Deterrence. For example, the lower gray curve is the buyer's value function when $s = Q^{-1}(\mu_0')$, therefore type-μ_0' buyer is indifferent between purchasing and learning. We let him break the tie by purchasing the item and he attains ex-ante trading surplus $V(\mu_0', Q^{-1}(\mu_0'))$. Varying μ_0', we can obtain the graph of buyer's ex-ante trading surplus $V(\cdot, Q^{-1}(\cdot))$, which is non-monotone in μ_0 and equals zero at $\underline{\mu}$ and $\overline{\mu}$ (see the orange dotted curve). Furthermore, since allocation is efficient and no learning cost is incurred, the joint surplus just equals the prior mean $\mathbb{E}(v|\mu_0)$, shown as the upper gray line. Then we can draw the seller's revenue $t^D(\mu_0)$ as the difference between the joint surplus and the buyer's ex-ante trading surplus, which is shown as the red curve.

Learning Deterrence is different from the extreme case in Corollary 1, since type-μ_0 buyer gives up his opportunistic information rent $V(\mu_0, Q^{-1}(\mu_0))$ which he could have been enjoyed if he continues to learn. The joint surplus is divided regard to this opportunistic information rent, which can be interpreted as the buyer's bargaining power. When the buyer is initially more uncertain, he enjoys larger benefit from learning and thereby attains larger bargaining power if the seller does not encourage learning. This gives a hint on the non-monotonicity of the buyer's expected trading surplus. Moreover, from Figure 2, $Q^{-1}(\mu_0)$ is also the highest s that is able to induce type-μ_0 buyer

[16] Figure 1 captures the same idea. The reader can decide which figure to look at.
[17] The buyer breaks indifference by purchasing the item immediately.

to learn,[18] which indicates that $V(\mu_0, Q^{-1}(\mu_0))$ is the highest opportunistic information rent that type-μ_0 buyer can ever get. Interestingly, it is attained by Learning Deterrence.

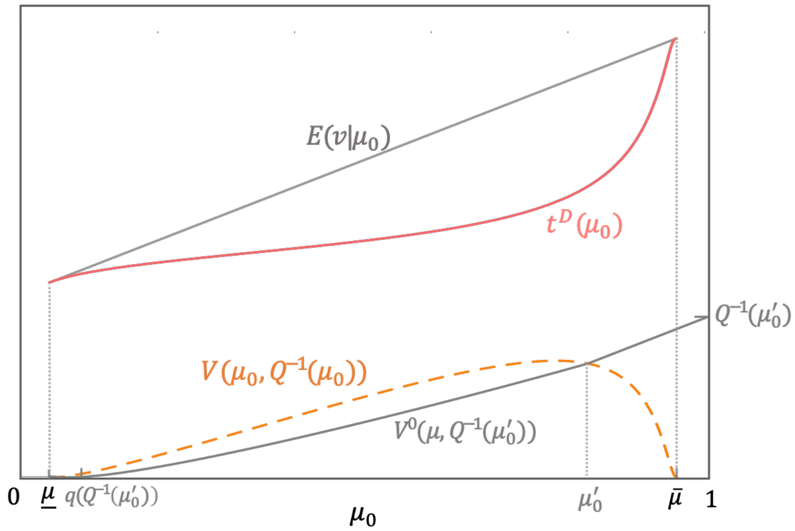

Figure 3: Allocation of trading surplus under Leaning Deterrence

Note that the olive-shaped region in Figure 2 captures the feasible set of stopping beliefs that an optimal mechanism could induce. Learning Deterrence—as the upper boundary—is a corner solution. Nevertheless, the seller can induce much more flexible buyer-learning. For the example, consider μ_0 denoted as in Figure 2, the gray region depicts the set of beliefs that the seller can enforce this type-μ_0 buyer to stop. In other words, if the seller is willing to encourage learning, the gray region is the set of buyer-learning that she can choose. In the next section, we study learning encouragement. In particular, we discuss the revenue-maximizing buyer-learning that the seller would like to enforce.

3 Encourage learning

To study the seller-optimal buyer-learning, we first characterize the selling mechanism with each element as a function of the buyer's stopping belief, so that such stopping belief is the buyer's best response.

[18] Suppose the buyer chooses to learn when he feels indifferent.

Lemma 2. *For a fixed $s \in [\underline{s}, \overline{s}]$, the return policy $(x_r(\mu, s), t_r(\mu, s))$ induces the buyer to stop learning at $\mu \in [q(s), Q(s)]$, where*

$$x_r(\mu, s) = \frac{V_1(\mu, s)}{v_h - v_l} \tag{6}$$

$$t_r(\mu, s) = \mathbb{E}(v|\mu)\frac{V_1(\mu, s)}{v_h - v_l} - V(\mu, s) \tag{7}$$

and $V_1(\mu, s)$ represents the partial derivative w.r.t μ. Furthermore, $t_r(\mu, s)$ is increasing in both μ and s, and the cross derivative is 0. $x_r(\mu, s)$ is increasing in μ.

Imposing the buyer's optimality (commonly known as smooth-pasting and value-matching) of stopping at μ, we obtain the explicit formula of the return policy. In particular, in Figure 4, the buyer obtains expected utility $\mathbb{E}(v|\mu)x_r - t_r$ if he requests a return, while he can attain continuation value $V(\mu, s)$ if he keeps learning. Making the buyer's return payoff tangent to his continuation value at μ, the buyer is willing to stop learning at μ and request such a return.[19] Therefore, by setting the proper return policy, the seller can enforce the buyer to stop learning even when there is positive information value.

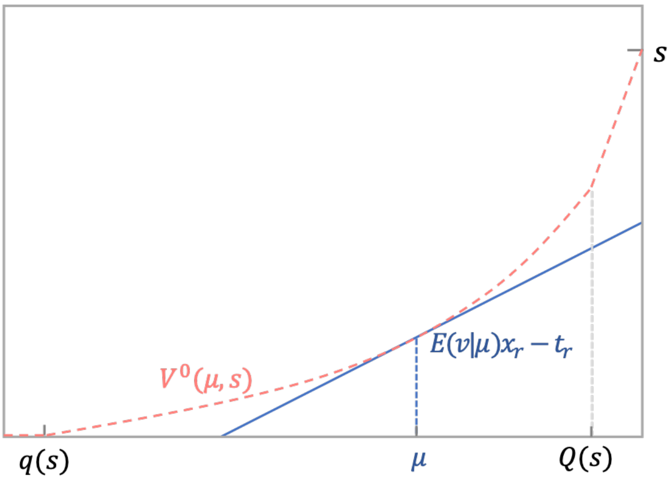

Figure 4: Partial learning with Stochastic Return

Interestingly, the seller can actually obtains larger return transfer if she enforces the buyer to stop earlier, even though the buyer enjoys larger opportunistic information

[19]We let the buyer break indifference in terms of the seller's favor. The reader can interpret this as a limit result. For example, the seller can decrease t_r by an arbitrarily small amount and meanwhile adjust x_r properly, so that the buyer strictly prefers to request a return at the same stopping belief.

rent. It is because the seller uses both x_r and t_r to compensate the buyer's opportunistic information rent. In particular, when the seller intends to induce larger stopping belief, she allows the buyer to keep the item with greater probability upon return, therefore the buyer is willing to make larger return transfer. With similar reasoning, the return transfer is increasing in s.

We define *full learning* as the situation where the buyer's continuation value for keep learning is zero, i.e., the stopping belief $\mu = q(s)$. While *partial learning* refers to the case where the buyer obtains strictly positive continuation value when he decides to return, i.e., the stopping belief $\mu > q(s)$. Hence, for the same s, Free Return induces full learning as the buyer stops learning at $q(s)$, while partial learning is induced by Stochastic Return with strictly positive $x_r > 0$.

$$\lim_{\mu \to q(s)} x_r(\mu, s) = 0 \qquad \text{and} \qquad \lim_{\mu \to q(s)} t_r(\mu, s) = 0$$

$$\lim_{\mu \to Q(s)} x_r(\mu, s) < 1 \qquad \text{and} \qquad \lim_{\mu \to Q(s)} t_r(\mu, s) < v_h - s$$

That is, fixing a price, Free Return $\{v_h - s, (0,0)\}$ is the left limit of Stochastic Return. Whereas, the right limit of Stochastic Return is strictly dominated by Learning Deterrence in terms of the seller's revenue.[20] Essentially, to induce the buyer to stop learning at $Q(s)$ with Stochastic Return, the joint surplus is smaller than $\mathbb{E}(v|\mu)$ and the buyer obtains the same trading surplus as in Learning Deterrence, therefore the seller is worse off.

Now we can formulate the seller's optimization problem for encouraging learning.

$$\max_{s \in [q^{-1}(\mu_0), Q^{-1}(\mu_0)]} \quad \max_{\mu} \quad \Pi(\mu, s) = \frac{\mu_0 - \mu}{1 - \mu}(v_h - s) + \frac{1 - \mu_0}{1 - \mu} t_r(\mu, s) \tag{$\mathcal{P}$}$$

$$\text{s.t.} \quad q(s) \leq \mu \leq Q(s)$$

$$\mu \leq \mu_0$$

The seller's expected revenue is a weighted average between the selling price and the return transfer. The weights depend on the buyer's stopping beliefs and the prior belief. Consider Figure 2 again, for a fixed μ_0, the seller chooses $s \in [q^{-1}(\mu_0), Q^{-1}(\mu_0)]$ to encourage learning. After which, she optimizes the expected revenue over the set of inducible return beliefs $\mu \in [q(s), Q(s)]$. Meanwhile, since the buyer's posterior belief decreases if no good news arrives, therefore the return belief $\mu \leq \mu_0$.

[20] $t_r(Q(s), s) < v_h - s$ implies $t_r(\mu, Q^{-1}(\mu)) < v_h - Q^{-1}(\mu) = t^D(\mu)$.

To study ($\mathcal{P}$), we go through two steps. Step 1, we study the internal maximization such that the seller chooses the optimal return beliefs while fixing s, denoted as ($\mathcal{R}$). Step 2, we study the external maximization along the solution path of ($\mathcal{R}$) and meanwhile reimpose the second constraint that we have ignored in ($\mathcal{R}$).

$$\max_{\mu \in [q(s), Q(s)]} \quad \Pi(\mu, s) \tag{$\mathcal{R}$}$$

Denote $\mu^*(s)$ as the local maximizer of ($\mathcal{R}$). In particular,

$$\mu^*(s) = \{\mu \in [q(s), Q(s)] : \Pi_1(\mu, s) = 0 \text{ and } \Pi_{11}(\mu, s) \leq 0\}.$$

Denote $\Pr(\text{return}) = \frac{1-\mu_0}{1-\mu}$ as the ex-ante chance of return. Rearranging the first order condition, we can evaluate the seller's trade-off among choosing the optimal stopping belief.

$$\underbrace{\Pr(\text{return})\frac{\partial t_r(\mu, s)}{\partial \mu}}_{\text{larger return transfer}} = \underbrace{[v_h - s - t_r(\mu, s)]\frac{d\Pr(\text{return})}{d\mu}}_{\text{more frequent return}}$$

That is, to increase the return belief, the seller can gain larger return transfer. But at the meantime, there is a larger chance of return and thereby the seller gets return transfer instead of the selling price more often. In Figure 5, while fixing s, we can draw the curve of $t_r(\cdot, s)$ within the domain $[q(s), Q(s)]$. $\mu^*(s)$ then depicts the locally optimal stopping belief, at which, the marginal gain from larger return transfer equals the marginal loss from more frequent return. Graphically, the blue line that connects $v_h - s$ and $t_r(\mu^*(s), s)$ is tangent to $t_r(\cdot, s)$. This also indicates that $\mu^*(s)$ is irrelevant to the prior belief μ_0.

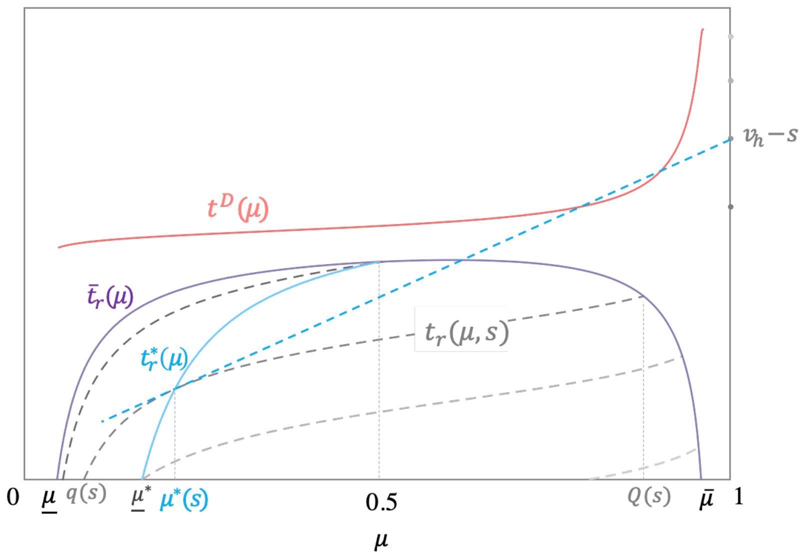

Figure 5: Optimal partial learning for fixed selling price

Note that $\bar{t}_r(\mu) := t_r(\mu, Q^{-1}(\mu))$ is the envelop of all inducible return transfers. To see how the locally optimal stopping belief changes with s. While increasing s, the selling price is reduced and the return transfer becomes larger (the gray dashed curve in Figure 5 shifts up), therefore the refund $v_h - s - t_r(\mu, s)$ that the seller commits to pay becomes smaller. The seller then cares less about the return rate and his incentive to gain larger safe revenue becomes relatively substantial. Thus, she optimally adapts to enforce a larger stopping belief—$\mu^*(s)$ is increasing in s. Furthermore, if s becomes sufficiently high,[21] the refund turns out to be sufficiently small and thereby to gain larger return transfer becomes the seller's dominant incentive. Hence she prefers to enforce stopping belief as large as possible and the upper boundary point $Q(s)$ is the optimal stopping belief. Conversely, if s is sufficiently small,[22] the dominant incentive is to reduce return rate and the seller prefers to induce stopping belief as small as possible, which makes the lower boundary point $q(s)$ the optimal stopping belief. Hence, for extreme s, $\mu^*(s)$ does not exist.

Lemma 3. *$\mu^*(s)$ is increasing in s. The range of $\mu^*(s)$ is $[\underline{\mu}^*, 0.5]$, where $\underline{\mu}^*$ is the solution of $\Pi_1(\mu, q^{-1}(\mu)) = 0$.*

Let $s^*(\mu)$ be the inverse function of $\mu^*(s)$. It specifies the amount of net surplus, under which the seller optimally induces a return belief μ. Then $t_r^*(\mu) := t_r(\mu, s^*(\mu))$ represents the return transfer that the seller obtains when μ is the locally optimal return belief. In Figure 5, $t_r^*(\mu)$ crosses the horizontal axis at $\underline{\mu}^*$ and crosses $\bar{t}_r(\mu)$ at 0.5 (shown as the blue curve), which is just the range of $\mu^*(s)$ in Lemma 3. Given this, $\Pi_1(\underline{\mu}^*, q^{-1}(\underline{\mu}^*)) = 0$ and $\Pi_1(0.5, Q^{-1}(0.5)) = 0$. The former one implies that when $s = q^{-1}(\underline{\mu}^*)$, inducing full learning is locally optimal for the seller. And the latter one implies that when $s = Q^{-1}(0.5)$, inducing no learning is locally optimal for the seller.[23] Hence, as long as $s \in (q^{-1}(\underline{\mu}^*), Q^{-1}(0.5))$, the optimal selling mechanism induces partial learning.

Given step 1, if the revenue-maximizing mechanism of $(\mathcal{P})$ turns out to be an interior Stochastic Return mechanism, it has to be located on the path of $t_r^*(\mu)$.[24] In this part, we want to show, compared to any Stochastic Return that induces partial learning, the seller is better off by *either* reducing s and enforcing full learning *or* raising s and

[21] In particular, $s > Q^{-1}(0.5)$.

[22] In particular, $s < q^{-1}(\underline{\mu}^*)$.

[23] Here, we refer to the Stochastic Return that induces stopping at 0.5.

[24] Once we pin down the amount of t_r and the return belief, we can pin down the exact selling mechanism.

enforcing no learning.

Theorem 1. *Stochastic return is dominated by either Learning Deterrence or Free Return.*

The formal proof is shown in the Appendix. Here we try to give an illustration of the main idea. Along the path of $s^*(\mu)$, the seller's expected revenue consists of two parts.

$$\Pi(\mu, s^*(\mu)) = t_r^*(\mu) + [1 - \Pr(\text{return})][v_h - s^*(\mu) - t_r^*(\mu)]$$

The first term is the seller's guaranteed revenue. While the second term refers to the extra revenue (in expectation) she can obtain if the buyer discovers a good news. If the seller intends to raise the stopping belief, she benefits from larger safe revenue, but suffers a loss by earning less extra revenue. In terms of the loss, there are two driven forces. 1) Larger stopping belief indicates shorter learning. Thereby the expected chance of obtaining a good news, $1 - \Pr(\text{return})$, is smaller; 2) By raising up the stopping belief, the seller endogenously decreases the magnitude of the extra revenue, $v_h - s^*(\mu) - t_r^*(\mu)$. Substitute $\Pi_1(\mu, s^*(\mu)) = 0$ into the seller's expected revenue, we can simplify the above expression.[25]

$$\Pi(\mu, s^*(\mu)) = t_r^*(\mu) + \frac{\partial t_r(\mu, s^*(\mu))}{\partial \mu}(\mu_0 - \mu)$$

The prior belief μ_0 only matters in terms of the loss. That is, for a fixed stopping belief $\mu \in [\underline{\mu}^*, 0.5]$, the marginal increment of safe revenue remains the same when varying the prior belief, however, the marginal loss gets larger when the prior is larger. In particular, there exists a cutoff prior such that if μ_0 goes above it, the dominant incentive is to reduce loss, which means that the seller is willing to decrease stopping belief. Conversely, if μ_0 is smaller than that cutoff, the dominant incentive is to increase the safe revenue and thereby the seller adjusts a higher stopping belief. It turns out that the cutoff prior is increasing in μ. Therefore, inversely, if we fix a prior belief, there exists a cutoff stopping belief.[26] Below which, the seller optimally encourages full learning, i.e., bringing down the stopping belief to $\underline{\mu}^*$. Above which, the seller optimally induces no learning, i.e., inducing instant stopping at the prior belief. As we mentioned before, induce no learning via Stochastic Return is strictly dominated

[25]This can be easily seen from Figure 5 where the tangent applies.

[26]If the cutoff stopping belief goes beyond $[\underline{\mu}^*, 0.5]$, then one incentive dominates the other everywhere. The optimal stopping belief is at the boundary.

by Learning Deterrence.[27] Therefore, Stochastic Return that induces partial learning is sub-optimal, the seller can attain larger expected revenue by either inducing full learning via Free Return or inducing no learning via Learning Deterrence.

Note that inducing stopping belief $\underline{\mu}^*$ with $s = q^{-1}(\underline{\mu}^*)$ corresponds to one particular Free Return mechanism, $\{v_h - q^{-1}(\underline{\mu}^*), (0,0)\}$. With different price, Free Return can induce different stopping beliefs. Substitute $\mu = q(s)$ into ($\mathcal{P}$),[28] we can derive the optimal Free Return mechanism.

$$\max_{s} \quad \Pi(q(s), s) = \frac{\mu_0 - q(s)}{1 - q(s)}(v_h - s)$$
$$\text{s.t.} \quad q^{-1}(\mu_0) \leq s \leq Q^{-1}(\mu_0)$$

Let $s^F(\mu_0) = \arg\max \Pi(q(s), s)$ be the unconstrained maximizer and $\Pi^F(\mu_0)$ as the corresponding revenue.

Lemma 4. *$s^F(\mu_0)$ is decreasing in μ_0. Both the selling price and expected revenue of the optimal Free Return are increasing in μ_0.*

The seller can induce longer learning with a larger s, which raises the expected probability of realizing a good news and thereby increases the chance of successful sale. Nevertheless, the selling price is lower with a larger s. Thus, there is an interior $s^F(\mu_0)$ that maximizes seller's expected revenue. With a larger prior belief, the chance of successful sale is higher and the seller puts more weight on the selling price. Therefore she locally adjusts a lower s to increase the selling price.

4 Optimal selling mechanism

Since Stochastic Return is always sub-optimal, the optimal selling mechanism is either Free Return or Leaning Deterrence. Denote $F = \{\mu_0 : t^D(\mu_0) \leq \Pi^F(\mu_0)\}$.

Theorem 2. *$F \subset (v_l/v_h, \overline{\mu})$ and it is either an empty set or a closed interval. If $\mu_0 \in F$, the optimal selling mechanism is $\{v_h - s^F(\mu_0), (0,0)\}$; If $\mu_0 \in [\underline{\mu}, \overline{\mu}] \setminus F$, the optimal selling mechanism is Learning Deterrence $\{t^D(\mu_0), (1, t^D(\mu_0))\}$.*

[27] In this scenario, when $\mu_0 > 0.5$, the upper boundary of stopping belief that the seller can induce is $\mu^*(Q^{-1}(\mu_0)) < \mu_0$. If we ignore this boundary constraint, then the Stochastic Return that induces stopping at 0.5 generates expected revenue as a weighted average between $t_r(0.5, Q^{-1}(0.5))$ and $v_h - Q^{-1}(0.5) = t^D(0.5)$, which is clearly smaller than $t^D(0.5) < t^D(\mu_0)$.

[28] $q^{-1}(\mu_0) \leq s$ implies $q(s) \leq \mu_0$. Hence we can drop the second constraint.

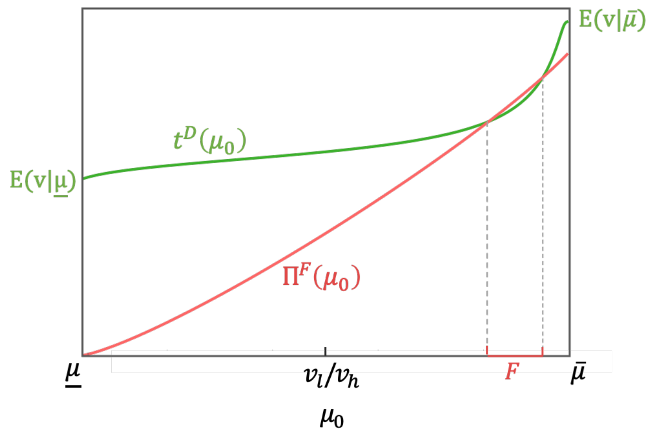

Figure 6: Expected Revenue of Free Return and Learning Deterrence

When learning is feasible to the buyer, the seller chooses the amount of buyer-learning she would like to induce optimally. Figure 6 depicts the expected revenue between Learning Deterrence and (the optimal) Free Return. The maximum of $\{t^D(\mu_0), \Pi^F(\mu_0)\}$ pins down the optimal selling mechanism. To interpret this Theorem 2, recall that $t^D(\mu_0) = \mathbb{E}(v|\mu_0) - V(\mu_0, Q^{-1}(\mu_0))$. If adopting Learning Deterrence, though allocation is efficient, the seller has to compensate the buyer with his opportunistic information rent. When μ_0 is sufficiently informative (equal to $\underline{\mu}$ or $\bar{\mu}$), it is obvious that the seller optimally deters learning since he can capture full joint surplus and compensating zero opportunistic information rent (Proposition 2). When the buyer's prior belief becomes more uncertain, the opportunistic information rent gets larger, to deter learning, the seller has to give away more allocation surplus to the buyer. Learning Deterrence turns to be less attractive. Moreover, Free Return wastes strictly positive amount of allocation surplus, but it allows the seller to charge a higher price compared to Learning Deterrence. When the nature endows the buyer a more optimistic prior, even Free Return can guarantee the seller a decent chance of successful sale. Hence the seller would optimally switch to Free Return if the prior belief implies a huge opportunistic information rent but is relatively optimistic.

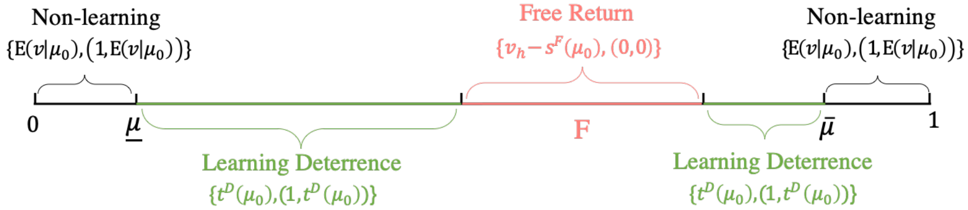

Figure 7: Optimal selling mechanism

Figure 7 completely characterizes the optimal selling mechanism when $\mu_0 \in [0, 1]$. When the buyer is initially informative enough and considers learning as sub-optimal anyway ($\mu_0 \notin [\underline{\mu}, \bar{\mu}]$), the seller captures the full allocation surplus and the buyer attains zero trading surplus. When learning becomes a valuable option to the buyer ($\mu_0 \in [\underline{\mu}, \bar{\mu}]$), the buyer obtains larger trading surplus if there is a larger opportunistic information rent. However, if the opportunistic information rent overwhelms the seller, the seller would instead encourage the buyer to learning. This drops the buyer's trading surplus by a strictly positive amount, since the joint allocation surplus is smaller but the seller obtains larger profit. Figure 8 depicts this decline.

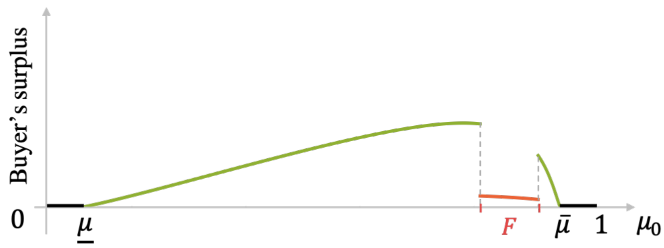

Figure 8: Buyer's trading surplus under optimal selling mechanism

Proposition 3. *Denote $\gamma = \frac{k}{\lambda}$. If $\gamma_1 < \gamma_2$, then $F(\gamma_2) \subseteq F(\gamma_1)$ and $[\underline{\mu}(\gamma_2), \overline{\mu}(\gamma_2)] \subset [\underline{\mu}(\gamma_1), \overline{\mu}(\gamma_1)]$. The seller's expected revenue from the optimal Free Return is decreasing in γ, while her revenue from Learning Deterrence is increasing in γ.*

We can interpret $\gamma = \frac{k}{\lambda}$ as a measure of learning efficiency. In particular, smaller γ indicates more efficient learning. It is intuitive that with more efficient learning, the buyer is more tempting to learn and thereby the opportunistic information rent that seller has to compensate to deter learning is larger. This implies a smaller revenue from Learning Deterrence. However, with Free Return, the seller can attain larger expected revenue as the buyer exerts longer learning and the ex-ante probability of successful sale becomes larger. Therefore, with a smaller γ, the set of priors that the seller optimally picks Free Return expands (see Figure 9 from the left to the right). Conversely, when γ is sufficiently high, the magnitude of opportunistic information rent is small enough and thereby Learning Deterrence dominates Free Return regardless of the prior belief, i.e., $F = \emptyset$.

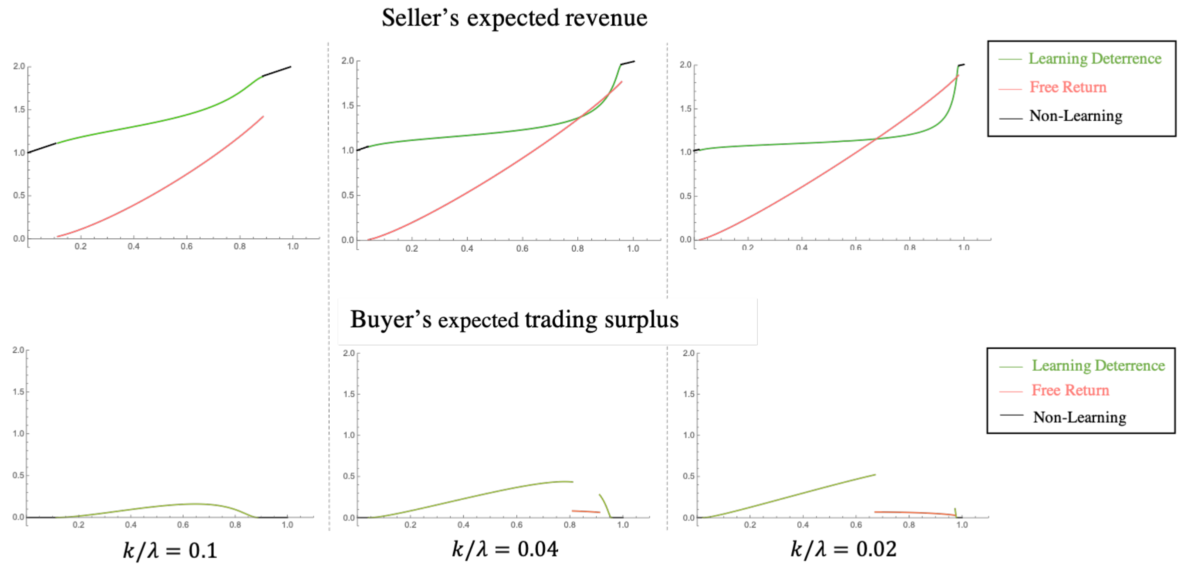

Figure 9: Comparative statics

Proposition 4. (A limit result)

$$\lim_{\gamma \to 0} \quad \max\{t^D(\mu_0), \Pi^F(\mu_0)\} = \begin{cases} v_l, & \mu_0 < \frac{v_l}{v_h} \\ \mu_0 v_h, & \mu_0 \geq \frac{v_l}{v_h} \end{cases}$$

Recall the well known result of selling a single indivisible good, where the buyer has private value $v \in \{v_h, v_l\}$ and the seller attaches a prior belief μ_0 as the probability of high value buyer. The revenue-maximizing mechanism suggests that, when $\mu_0 < \frac{v_l}{v_h}$, setting a price equal v_l (known as mass market); when $\mu_0 \geq \frac{v_l}{v_h}$, setting a price equal v_h (known as niche market). Intuitively, when the seller believes that there is a larger chance to match with a high value buyer, he is willing to charge a high price and sacrifice the trading opportunity with the low value buyer, vise versa. In our setup, when γ goes to zero, the seller's revenue converges to the revenue he can gets with a privately perfect-informed buyer. This is because, if γ converges to zero, then the buyer can learn almost perfect information, therefore with Free Return, the seller optimally sets the selling price arbitrarily close to v_h and lets go the buyer almost sure to have low value. While with Leaning Deterrence, the seller has to set the price arbitrarily close to v_l, otherwise, the buyer always has incentive to learn to avoid consuming the item when his true value is low. The ratio $\frac{v_l}{v_h}$ pins down the cutoff of prior belief that the seller is indifferent between Free Return and Learning Deterrence.

5 Conclusion

This paper discusses the optimal mechanism when the buyer can privately acquire information about a product. The seller designs the mechanism to influence the buyer's learning value and thereby indirectly controls the buyer's learning behaviour. We find that the optimal selling mechanism either induces full learning or deters the buyer's private learning. And the buyer obtains different trading surplus when his prior belief varies. However, our results rely on the assumption that buyer's prior belief is common information. Therefore, it naturally brings out the next question: what if the buyer has private information *a priori*? Then clearly he has incentive to misreport his private information. If the seller can design a time-contingent return policy, is it possible for her to use it to screen the buyers endowed with different prior beliefs? We believe this is an interesting follow-up question.

References

Armstrong, Mark and Jidong Zhou, " Search Deterrence," *The Review of Economic Studies*, 07 2015, *83* (1), 26–57.

Bergemann, Dirk and Juuso Valimaki, "Learning and Strategic Pricing," *Econometrica*, 1996, *64* (5), 1125–1149.

_ **and** _ , "Experimentation in Markets," *The Review of Economic Studies*, 2000, *67* (2), 213–234.

_ **and Martin Pesendorfer**, "Information structures in optimal auctions," *Journal of Economic Theory*, 2007, *137* (1), 580–609.

Bonatti, Alessandro, "Menu Pricing and Learning," *American Economic Journal: Microeconomics*, August 2011, *3* (3), 124–63.

Guo, Yingni, Hao Li, and Xianwen Shi, "Optimal Discriminatory Disclosure," *Working paper*, 2020.

Johnson, Justin P. and David P. Myatt, "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review*, June 2006, *96* (3), 756–784.

Keller, Godfrey, Sven Rady, and Martin Cripps, "Strategic Experimentation with Exponential Bandits," *Econometrica*, 2005, *73* (1), 39–68.

Lang, Ruitian, "Try before you buy: A theory of dynamic information acquisition," *Journal of Economic Theory*, 2019, *183*, 1057–1093.

Li, Hao and Xianwen Shi, "Discriminatory information disclosure," *American Economic Review*, 2017, *107* (11), 3363–85.

Mensch, Jeffrey, "Screening inattentive agents," *Working paper*, 2020.

Pease, Marilyn, "Shopping for Information: The Implications of Consumer Learning for Optimal Pricing and Product Design," *working paper*, 2018.

Pomatto, Luciano, Philipp Strack, and Omer Tamuz, "The cost of information," *Working paper*, 2019.

Ravid, Doron, Anne-Katrin Roesler, and Balázs Szentes, "Learning before trading: on the inefficiency of ignoring free information," *working paper*, 2019.

Roesler, Anne-Katrin and Balázs Szentes, "Buyer-optimal learning and monopoly pricing," *American Economic Review*, 2017, *107* (7), 2072–80.

Shi, Xianwen, "Optimal auctions with information acquisition," *Games and Economic Behavior*, 2012, *74* (2), 666–686.

Appendix

Proposition 1

Proof. We prove this proposition by introducing a claim first. Recall that

$$Q(s) = \{\mu : V(\mu, s) = \mathbb{E}(v|\mu) - (v_h - s)\}.$$

Hence, by setting $s = D(\mu) := Q^{-1}(\mu)$, the type-$\mu$ buyer is indifferent between accepting the price or exerting learning and $\tilde{\mu}(\mu)$ is the optimal quitting belief.

Claim 1. *Let $\tilde{\mu}(\mu) := q(D(\mu))$ with domain $[\underline{\mu}, \overline{\mu}]$. $\tilde{\mu}(\mu) \leq \mu$ and the equality holds only at the two end points. $\tilde{\mu}(\mu)$ is increasing and symmetric about the line 1-μ. $\mu - \tilde{\mu}(\mu)$ is increasing and then decreasing.*

Proof. Substitute $D(\mu)$ into $V(\mu, s) = \mathbb{E}(v|\mu) - (v_h - s)$, we have,

$$V(\mu, D(\mu)) = \mathbb{E}(v|\mu) - (v_h - D(\mu)) \tag{8}$$

By implicit differentiation w.r.t. μ, we have,

$$\frac{dD(\mu)}{d\mu} = \frac{k(k - \lambda D(\mu))}{\lambda^2(1-\mu)^2\mu D(\mu)} = \frac{k[\tilde{\mu} - 1]}{\lambda(1-\mu)^2\mu} < 0 \tag{9}$$

Besides,

$$\frac{d\tilde{\mu}}{d\mu} = \frac{d[k/\lambda D(\mu)]}{d\mu} = \frac{k^2(\lambda D(\mu) - k)}{\lambda^3(1-\mu)^2\mu D(\mu)^3} = \frac{\tilde{\mu}^2(1-\tilde{\mu})}{(1-\mu)^2\mu}$$

This is a differential equation with the initial point $(\underline{\mu}, \underline{\mu})$.[29] The solution is,[30]

$$-\frac{1}{\tilde{\mu}} - \log[1-\tilde{\mu}] + \log[\tilde{\mu}] = \frac{1}{1-\mu} - \log[1-\mu] + \log[\mu] - \frac{\lambda(v_h - v_l)}{k} \tag{10}$$

Denote the LHS as $f(\tilde{\mu})$ and the RHS as $g(\mu)$. The domain of both functions is $[\underline{\mu}, \overline{\mu}]$ and $f(\cdot) = g(\cdot)$ at the two end points. Note that $f'(\cdot) > g'(\cdot)$ when the both arguments are smaller then 0.5 and $f'(\cdot) < g'(\cdot)$ when both arguments are larger then

[29] To verify $(\underline{\mu}, \underline{\mu})$ is an initial point. Recall that $\underline{\mu} = q(\bar{s})$. The binding Learning-Feasibility constraint implies that $\mathbb{E}(v|\underline{\mu}) - (v_h - \bar{s}) = 0$. Meanwhile $V(q(\bar{s})), \bar{s}) = V(\underline{\mu}, \bar{s}) = 0$. Substitute $\underline{\mu}$ into equation (8), $V(\underline{\mu}, D(\underline{\mu})) = \mathbb{E}(v|\underline{\mu}) - (v_h - D(\underline{\mu}))$, therefore $D(\underline{\mu}) = \bar{s}$. Then $\tilde{\mu}(\underline{\mu}) = \frac{k}{\lambda D(\underline{\mu})} = \underline{\mu}$.

[30] The general solution is $-\frac{1}{\tilde{\mu}} - \log[1-\tilde{\mu}] + \log[\tilde{\mu}] = \frac{1}{1-\mu} - \log[1-\mu] + \log[\mu] + C$. Conditional on the initial point, $(\underline{\mu}, \underline{\mu})$, we can solve $C = -\frac{\lambda(v_h - v_l)}{k}$. Same result holds if we take $(\overline{\mu}, \overline{\mu})$ as the initial point.

0.5.[31] Therefore $f(\cdot)$ and $g(\cdot)$ cross only at the two boundary points and therefore $\tilde{\mu}(\mu) < \mu$ for all $\mu \in (\underline{\mu}, \overline{\mu})$. For $\tilde{\mu}(\mu)$ to be symmetric about $1-\mu$, note that the reflection point of $(\mu, \tilde{\mu})$ over line $1-\mu$ is $(1-\tilde{\mu}, 1-\mu)$. And it is easy to check that for any point $(\mu, \tilde{\mu})$ that equation (10) holds, by substituting $(1-\tilde{\mu}, 1-\mu)$, equation (10) still holds. Now we show that $\mu - \tilde{\mu}(\mu)$ is single-peaked in μ and it increases first and then decreases. Note that $\tilde{\mu}'(\underline{\mu}) < 1$ and $\tilde{\mu}'(\overline{\mu}) > 1$, therefore if $\tilde{\mu}'(\mu) = 1$ has a unique solution, then we are done. To show this, $\frac{d\tilde{\mu}}{d\mu} = \frac{\tilde{\mu}^2(1-\tilde{\mu})}{(1-\mu)^2\mu} = 1$ implies $\tilde{\mu}(\mu) = 1-\mu$.[32] Since $\tilde{\mu}(\mu)$ is monotone in μ and symmetric about $1-\mu$. Hence, $\tilde{\mu}'(\mu) = 1$ has a unique interior solution. $\square$

Note that $Q(D(\mu)) - q(D(\mu)) = \mu - \tilde{\mu}(\mu)$. Take derivative w.r.t. μ, $(Q' - q')D' = 1 - \tilde{\mu}'$. Since $D'(\mu) < 0$, $Q'(s) - q'(s)$ is positive first and then negative. Hence $Q(s) \geq q(s)$ and the difference is single-peaked in s and $Q(s) = q(s)$ at $\underline{s}$ and $\bar{s}$. Therefore the Learning-Feasibility constraint holds when $s \in [\underline{s}, \bar{s}]$,[33] and the construction in Proposition 2 (2) is optimal based on the standard arguments in exponential bandits problem. In terms of (1), since the Learning-Feasibility constraint fails when $s \notin [\underline{s}, \bar{s}]$ and the buyer is still willing to purchase at the quitting belief. Therefore no learning is optimal. $\square$

Lemma 1

Proof. To simplify exposition, we omit the notion of s in the buyer's value function since the lemma is true for any fixed price. Let $V_B(\mu(\tau))$ be the buyer's value function of performing before-trading learning and it is characterized by the Bellman equation below.

$$V_B(\mu(\tau)) = \max\{0, V_P(\mu(\tau)), -kd\tau + \mu(\tau)\lambda d\tau s + (1-\mu(\tau)\lambda d\tau)V_B(\mu(\tau + d\tau))\} \quad (11)$$

At time τ, if the buyer continues to learn for a instant of time $d\tau$, with probability $\mu(\tau)\lambda d\tau$, a Poisson jump arrives and he purchases the item obtaining net surplus s; with the remaining probability, no news arrives and his belief moves to $\mu(\tau + d\tau)$ and the corresponding value is $V_B(\mu(\tau + d\tau))$. The buyer can also stop before-trading learning and choose the outside option to get payoff 0, or switch to post-trading learning by purchasing the item and obtaining purchase value $V_P(\mu(\tau))$, which

[31] $f' = \frac{1}{\tilde{\mu}^2 - \tilde{\mu}^3}$ and $g' = \frac{1}{(1-\mu)^2\mu}$.

[32] $\tilde{\mu}^2(1-\tilde{\mu}) = (1-\mu)^2\mu$ could have three solutions, $\tilde{\mu} = \mu = 0$, $\tilde{\mu} = \mu = 1$ or $\tilde{\mu} = 1-\mu$. The previous two cannot be true when $\mu \in [\underline{\mu}, \overline{\mu}]$.

[33] Note that if $\lim_{\mu \to 1} V(\mu, s) = s - \frac{k}{\lambda} < s$. Hence $\mathbb{E}(v|\cdot) - (v_h - s)$ crosses $V(\cdot, s)$ from below if there exists a $Q(s) > q(s)$. Given the convexity of $V(\cdot, s)$, the Learning-Feasibility constraint holds.

is characterized below in an analogous manner.

$$
\begin{aligned}
V_P(\mu(\tau)) = \max\{&\mathbb{E}(v|\mu(\tau)) - t_b, \mathbb{E}(v|\mu(\tau))x_r - t_r, \\
&- kd\tau + \mu(\tau)\lambda d\tau s + (1 - \mu(\tau)\lambda d\tau)V_P(\mu(\tau + d\tau))\}
\end{aligned}
\tag{12}
$$

Note that while the buyer purchases the item, he instantaneously abandons his outside option. In other words, upon stopping, he can either consume the item or return it according to the pre-specified return policy. Conditional on learning, the three Bellman equations (11), (12) and (2) lead to the same differential equation (ODE). Meanwhile, it is obvious that $V_B(\mu) \geq V_P(\mu)$ and $V_B(\mu) \geq V^0(\mu)$.

To show this Lemma, we prove the following equality.

$$
\max\{V_B(\mu), V_P(\mu)\} = V^0(\mu), \ \forall \mu. \tag{13}
$$

Suppose the seller intends to set a predatory return policy, with which if the buyer purchases the item and stops post-trading learning at belief μ,[34] he obtains payoff $V_P(\mu) < V_B(\mu)$. Then a rational buyer would instead not purchase the item and perform learning before-trading merely, which implies $V_B(\mu) = V^0(\mu)$ and the above equality holds.

Moreover, suppose the seller instead offers a benevolent return policy intending to reward the buyer from purchasing the item. With which, the buyer purchases the item at some point and while he stops post-trading learning at belief μ and requests a return, he gets payoff $V_P(\mu) > V^0(\mu)$. We can then calculate the return transfer which equals the trading surplus minus the buyer's payoff.[35]

$$
t_r = \mathbb{E}(v|\mu)x_r - V_P(\mu) = \mathbb{E}(v|\mu)\frac{V_P'(\mu)}{v_h - v_l} - V_P(\mu)
$$

The (ODE) is a general solution of $V_P(\mu)$. Hence,

$$
(1 - \mu)\mu\lambda V_P'(\mu) + \mu\lambda V_P(\mu) = \mu\lambda s - k
$$

The slope $V_P'(\mu)$ and the magnitude $V_P(\mu)$ of the buyer's continuation value are the substitutes that the seller can adjust to enforce the same stopping belief. For the purpose of maximizing profit, the seller will reduce $V_P(\mu)$ and raise $V_P'(\mu)$ (constrained by the above differential equation) to increase the return transfer and at the mean time

[34]Since the buyer always stops learning if his belief jumps to 1, to simplify exposition, when we say stopping belief, we mean the non-degenerate stopping belief.

[35]For μ to be an incentive compatible stopping belief, the smooth pasting condition should be satisfied, $V_P'(\mu) = d[\mathbb{E}(v|\mu)x_r - t_r]/d\mu = (v_h - v_l)x_r$. This implies the second equality.

preserve the same buyer's optimal stopping,[36] which is a contradiction of optimality and we obtain the condition (13). □

Corollary 1

Proof. Given Lemma 1, for all optimal selling mechanisms, $V_B(\mu_0, s) = V^0(\mu_0, s)$. Endowed with sufficiently pessimistic or optimistic prior, the buyer obtains payoff $V_B(\mu_0, v_h - t_b) = \max\{0, \mathbb{E}(v|\mu) - t_b\}$. Suppose the seller offers a "No Return" mechanism with price $t_b = \mathbb{E}(v|\mu)$. Then the buyer weakly prefers to purchase the item and the seller can extract all trading surplus. To verify that this selling mechanism $\{\mathbb{E}(v|\mu_0), (1, \mathbb{E}(v|\mu_0))\}$ is optimal, note that the seller's profit equals to the total trading surplus minus the surplus that buyer obtains from the trade. In this case, the trading surplus is the largest since trading happens with probability one and the posterior mean of true value equals to the prior mean for all possible learning dynamic, meanwhile the buyer gets zero trading surplus. □

Proposition 2

Proof. Notice that when $s \geq D(\mu_0)$, Lemma 1 and Proposition 1 imply that

$$V_B(\mu_0, s) = \mathbb{E}(v|\mu_0) - (v_h - s).$$

Therefore trading happens with probability one. The seller's revenue just equals the selling price $v_h - s$ and it is optimal to set $t_b = v_h - D(\mu_0)$.

First, we prove (1). At $\underline{\mu}$ and $\overline{\mu}$, the Learning-Feasibility constraint binds. This implies $V(\underline{\mu}, D(\underline{\mu})) = V(\underline{\mu}, Q^{-1}(\underline{\mu})) = 0$ and $V(\overline{\mu}, D(\overline{\mu})) = V(\overline{\mu}, Q^{-1}(\overline{\mu})) = 0$. Rearrange equation (8),

$$V(\mu, D(\mu)) = D(\mu) - (1 - \mu)(v_h - v_l)$$

Take derivative w.r.t μ,

$$\frac{dV(\mu, D(\mu))}{d\mu} = (v_h - v_l)\left[-A\frac{(1-\tilde{\mu})}{(1-\mu)^2\mu} + 1\right]$$

where $A = \frac{k}{\lambda(v_h - v_l)} = (1 - \underline{\mu})\underline{\mu} \in (0, \frac{1}{4})$.[37] It is easy to verify $\frac{dV(\mu, D(\mu))}{d\mu} = 0$ at $\underline{\mu}$ or $\overline{\mu}$. Since $V(\mu, D(\mu)) > 0$ when $\mu \in (\underline{\mu}, \overline{\mu})$, to prove that $V(\mu, D(\mu))$ is single-peaked

[36] By inducing the same stopping beliefs (including 1), the ex-ante probabilities of return and successful sale are the same regardless of when the buyer switches to post-trading.

[37] From the binding Learning-Feasibility constraint, we can get $\frac{k}{\lambda(v_h - v_l)} = (1 - \underline{\mu})\underline{\mu} = (1 - \overline{\mu})\overline{\mu}$. Therefore, $\underline{\mu} = 1 - \overline{\mu} \in (0, 0.5)$. Hence $A \in (0, \frac{1}{4})$.

in μ, we only need to show that $\frac{dV(\mu,D(\mu))}{d\mu} = 0$ has a unique solution when $\mu \in (\underline{\mu}, \overline{\mu})$. That is, the two equations below have a unique solution when $\mu \in (\underline{\mu}, \overline{\mu})$, since $\tilde{\mu}$ is the implicit solution of (10).

$$-A\frac{(1-\tilde{\mu})}{(1-\mu)^2\mu} + 1 = 0 \tag{14}$$

$$-\frac{1}{\tilde{\mu}} + \log\left[\frac{\tilde{\mu}}{1-\tilde{\mu}}\right] = \frac{1}{1-\mu} + \log\left[\frac{\mu}{1-\mu}\right] - \frac{1}{A} \tag{15}$$

Substituting equation (14) into (15), we have,

$$h(\mu) := -\frac{A}{A-(1-\mu)^2\mu} + \log\left[\frac{A-(1-\mu)^2\mu}{(1-\mu)^2\mu}\right] - \left(\frac{1}{1-\mu} + \log\left[\frac{\mu}{1-\mu}\right] - \frac{1}{A}\right) = 0$$

Now we want to show that $h(\mu) = 0$ has a unique solution for $\mu \in (\underline{\mu}, \overline{\mu})$. In particular, as we can verify that $h(\mu) = 0$ at $\underline{\mu}$ and $\overline{\mu}$, we want to show $h(\mu)$ first decreases and then increases and then decreases on $[\underline{\mu}, \overline{\mu}]$.

$$h'(\mu) = \frac{1}{(1-\mu)^2\mu}\left[\frac{y(\mu)}{z(\mu)} - 1\right]$$

where $y(\mu) := A^2(3\mu-1)(1-\mu)$ and $z(\mu) := [A-(1-\mu)^2\mu]^2$. $y(\mu)$ is a second order polynomial function that is negative when $\mu < 1/3$, and it increases on μ if $\mu < 2/3$ and decreases on μ if $\mu > 2/3$. $z(\mu)$ is a high order polynomial function and $z'(\mu) = 0$ has a most 4 roots: $1/3, 1$ and the roots of $(1-\mu)^2\mu - A = 0$ (at most two roots).[38] We can show that $z(\mu)$ crosses $y(\mu)$ in the support $[\underline{\mu}, \overline{\mu}]$ twice, first from above and then from below.[39] Done.

Next, the monotonicity of $t^D(\mu_0) = v_h - D(\mu_0)$ can be directly obtained from (9). Moreover $t^D(\mu_0) = \mathbb{E}(v|\mu_0) - V(\mu_0, D(\mu_0))$ and $V(\underline{\mu}, D(\underline{\mu})) = V(\overline{\mu}, D(\overline{\mu})) = 0$, therefore $t^D(\underline{\mu}) = \mathbb{E}(v|\underline{\mu})$ and $t^D(\overline{\mu}) = \mathbb{E}(v|\overline{\mu})$. □

Lemma 2

[38] $z'(\mu) = 2[(1-\mu)^2\mu - A](3\mu-1)(\mu-1)$. The derivative of $(1-\mu)^2\mu - A$ is $(3\mu-1)(\mu-1)$. Hence $(1-\mu)^2\mu - A$ is increasing if $\mu < 1/3$ and decreasing afterwards. When $A < 4/27$, $(1-\mu)^2\mu - A = 0$ has two distinct roots, $r_1 < 1/3 < r_2$. When $A = 4/27$, there is a unique root $1/3$. When $A > 4/27$, there is no root. Regardless of A, $(1-\mu)^2\mu - A < 0$ when $\mu = \underline{\mu}, \overline{\mu}$.

[39] (1) Suppose $A < 4/27$, then $z(\mu) > y(\mu)$ for $\mu \leq 1/3$, $z(r_2) = 0 < y(r_2)$ and $z(\overline{\mu}) > y(\overline{\mu})$. Therefore $z(\mu)$ double crosses $y(\mu)$. (2) Suppose $A = 4/27$, then $z(\mu) > y(\mu)$ for $\mu < 1/3$, $z(1/3) = y(1/3)$, $z'(1/3) = 0 < y'(1/3)$ and $z(\overline{\mu}) > y(\overline{\mu})$. Therefore $z(\mu)$ double crosses $y(\mu)$. (3) Suppose $A \in (4/27, 1/4)$, then $z'(\mu) < 0$ when $\mu < 1/3$ and $z'(\mu) \geq 0$ when $\mu \geq 1/3$. We can check that $z(1/2) < y(1/2)$ for $A \in (4/27, 1/4)$, hence given $y(\underline{\mu}) < z(\underline{\mu})$ and $y(\overline{\mu}) < z(\overline{\mu})$, we have the same double crossing.

Proof. Given Lemma 1, $V_B(\cdot, s) = V^0(\cdot, s) \geq V_P(\cdot, s)$ on the domain $[0, 1]$. To induce the buyer to stop at belief μ different from $q(s)$, $V_P(\mu, s)$ must be equal to $V^0(\mu, s)$. Otherwise, the buyer strictly prefers to stay in before-trading learning and does not stop. Furthermore, to ensure that it is incentive compatible for the buyer to stop at belief μ and request the return (x_r, t_r), the buyer's expected payoff from requesting return, $\mathbb{E}(v|\cdot)x_r - t_r$, should smoothly pass $V^0(\cdot, s)$ at μ. Besides, the induced stopping belief μ must belong to the set $[q(s), Q(s)]$, in which $V^0(\mu, s) = V(\mu, s)$. Above all,

$$\text{value matching: } \mathbb{E}(v|\mu)x_r - t_r = V(\mu, s)$$

$$\text{smooth pasting: } \frac{d[\mathbb{E}(v|\mu)x_r - t_r]}{d\mu} = V_1(\mu, s)$$

We then obtain the expression of x_r and t_r. Specifically,

$$t_r(\mu, s) = -\frac{kv_l - \lambda\mu v_l s - k\mu v_h \left[\log\left(\frac{\mu}{1-\mu}\right) - \log\left(\frac{k}{\lambda s - k}\right)\right]}{\lambda\mu(v_h - v_l)}$$

Take partial derivative w.r.t μ and s separately,

$$\frac{\partial t_r(\mu, s)}{\partial \mu} = \frac{k\mathbb{E}(v|\mu)}{\lambda(1-\mu)\mu^2(v_h - v_l)} > 0$$

$$\frac{\partial t_r(\mu, s)}{\partial s} = \frac{\mathbb{E}(v|q(s))}{(1-q(s))(v_h - v_l)} > 0$$

Moreover, since $V(\cdot, s)$ is convex in μ, therefore $x_r(\cdot, s)$—proportional to $V_1(\cdot, s)$—is increasing in μ. $\square$

Lemma 3

Proof. Recall that $s^*(\mu)$ is the inverse function of $\mu^*(s)$ and $D(\mu) := Q^{-1}(\mu)$. Use implicit differentiation to $\Pi_1(\mu, s) = 0$, we have,

$$\frac{ds^*(\mu)}{d\mu} = \frac{k[(2\mu - 1)\mu v_h - 2(1-\mu)^2 v_l](k - \lambda s^*(\mu))}{\lambda^2(1-\mu)\mu^3 v_h s^*(\mu)}$$

Hence, it is easy to verify that $s^*(\mu)$ is increasing on μ when $\mu \leq 0.5$. Equivalently, $\mu^*(s)$ is increasing in s if $\mu^* \leq 0.5$. Next, we want to pin down the range of $\mu^*(s)$.

Claim 2. *The domain of $\bar{t}_r(\mu)$ is $[\underline{\mu}, \overline{\mu}]$. $\bar{t}_r(\mu) = 0$ at the two end points. It increases first and then decreases. $\bar{t}_r(\mu) = t_r^*(\mu)$ has a unique solution* 0.5.

Recall that $\bar{t}_r(\mu, Q^{-1}(\mu))$. It is obvious that $t_r(\underline{\mu}, Q^{-1}(\underline{\mu})) = t_r(\overline{\mu}, Q^{-1}(\overline{\mu})) = 0$. Take total derivative of $\bar{t}_r(\mu)$ w.r.t μ,

$$\frac{d\bar{t}_r(\mu)}{d\mu} = \frac{\partial t_r(\mu, D(\mu))}{\partial \mu} + \frac{\partial t_r(\mu, D(\mu))}{\partial s}\frac{dD(\mu)}{d\mu} = \frac{(1-\underline{\mu})\underline{\mu}}{(1-\mu)\mu}\left[\frac{\mathbb{E}(v|\mu)}{\mu} - \frac{\mathbb{E}(v|\tilde{\mu}(\mu))}{1-\mu}\right]$$

The term in the bracket is decreasing. It's positive when $\mu = \tilde{\mu}(\mu) = \underline{\mu}$ and it's negative when $\mu = \tilde{\mu}(\mu) = \overline{\mu}$. Hence $\bar{t}_r(\mu)$ is increasing first and then decreasing. Now we want to show $t_r^*(0.5) = \bar{t}_r(0.5)$. From equation (8), (6) and (7), we have,

$$\mathbb{E}(v|\mu) - (v_h - D(\mu)) = V(\mu, D(\mu)) = \mathbb{E}(v|\mu)x_r(\mu, D(\mu)) - t_r(\mu, D(\mu))$$

which implies,[40]

$$v_h - D(\mu) - t_r(\mu, D(\mu)) = \mathbb{E}(v|\mu)[1 - x_r(\mu, D(\mu))] = \frac{k\mathbb{E}(v|\mu)}{\lambda(1-\mu)\mu(v_h - v_l)} \tag{16}$$

Recall that $\frac{\partial t_r(\mu,s)}{\partial \mu} = \frac{k\mathbb{E}(v|\mu)}{\lambda(1-\mu)\mu^2(v_h-v_l)}$ and substitute it into $\Pi_1(\mu, s) = 0$, we have

$$v_h - s^*(\mu) - t_r(\mu, s^*(\mu)) = \frac{k\mathbb{E}(v|\mu)}{\lambda\mu^2(v_h - v_l)} \tag{17}$$

At the crossing point of $t_r(\mu, s^*(\mu))$ and $t_r(\mu, D(\mu))$, $s^*(\mu) = D(\mu)$. Therefore, we can then solve $\mu = 0.5$ as the unique solution that equation (16) and (17) equalize. *Q.E.D*

Note that,

$$\frac{d\bar{t}_r}{d\mu} = \frac{\partial t_r(\mu, D)}{\partial \mu} + \frac{\partial t_r(\mu, D)}{\partial s}\frac{dD}{d\mu}$$
$$\frac{dt_r^*}{d\mu} = \frac{\partial t_r(\mu, s^*)}{\partial \mu} + \frac{\partial t_r(\mu, s^*)}{\partial s}\frac{ds^*}{d\mu}$$

Since $\frac{\partial t_r(\mu,s)}{\partial \mu}$ is independent of s, $\frac{\partial t_r(\mu,D)}{\partial \mu} = \frac{\partial t_r(\mu,s^*)}{\partial \mu}$. Besides, when $\mu \leq 0.5$, $\frac{dD}{d\mu} < 0$ and $\frac{ds^*}{d\mu} > 0$. Hence, the slope of $\bar{t}_r$ is smaller than t_r^*. That is, if we reduce μ from 0.5, $t_r^*(\mu)$ decreases faster than $\bar{t}_r(\mu)$. Denote $\underline{\mu}^*$ as the solution that $t_r^*(\mu) = 0$. Obviously, $\underline{\mu}^* > \underline{\mu}$. To pin down $\underline{\mu}^*$, note that $t_r(\mu, s) = 0$ implies $s = q^{-1}(\mu)$. Hence, $t_r(\underline{\mu}^*, s^*(\underline{\mu}^*)) = 0$ implies $s^*(\underline{\mu}^*) = q^{-1}(\underline{\mu}^*)$. Then, we substitute it back to the first order condition, $\Pi_1(\underline{\mu}^*, q^{-1}(\underline{\mu}^*)) = 0$. That is, $\underline{\mu}^*$ is the solution of below.

$$\Pi_1(\mu, q^{-1}(\mu)) = \frac{(\mu_0 - 1)(\lambda\mu^2 v_h(v_h - v_l) - k(2\mu(v_h - v_l) + v_l))}{\lambda(1-\mu)^2\mu^2(v_h - v_l)} = 0$$

Explicitly, $\underline{\mu}^* = \frac{k}{\lambda v_h} + (\frac{k}{\lambda v_h}(\frac{k}{\lambda v_h} + \frac{v_l}{v_h - v_l}))^{\frac{1}{2}}$.[41] Therefore, $[\underline{\mu}^*, 0.5]$ is the domain of $t_r^*(\mu)$ which is also the range of $\mu^*(s)$.

[40]Where the second equality is directly obtained by substituting (ODE) and $\mathbb{E}(v|\mu) - (v_h - D(\mu)) = V(\mu, D(\mu))$ into equation (6).

[41]Since $\lambda\mu^2 v_h(v_h - v_l) - k(2\mu(v_h - v_l) + v_l)$ is increasing on $\mu > 0$ (the derivative of it is $-2(k - \lambda v_h \mu)(v_h - v_l) > 0$) and it is negative when μ is small and is positive when μ is large. Hence $\Pi_1(\mu, q^{-1}(\mu))$ single crosses 0 from above and $\underline{\mu}^*$ is unique.

Now we verify that $\Pi_{11}(\mu, s^*(\mu)) \leq 0$ on $[\underline{\mu}^*, 0.5]$.

$$\frac{\partial \Pi(\mu,s)}{\partial \mu} = \frac{(1-\mu_0)}{(1-\mu)^2(v_h - v_l)} \underbrace{\left[v_h(-v_h + s + v_l) + \frac{k(\mu(v_h - 2v_l) + v_l)}{\lambda \mu^2} + \frac{k v_h (\log[\frac{\mu}{1-\mu}] - \log[\frac{k}{\lambda s - k}])}{\lambda} \right]}_{\equiv \Upsilon(\mu)}$$

Note that $\frac{\partial \Pi(\mu,s)}{\partial \mu}|_{\mu=\mu^*(s)} = 0$, therefore $\Upsilon(\mu^*(s)) = 0$. Moreover, $\Upsilon(\mu)$ is decreasing in μ.[42] This implies that when the first order condition holds, the second order condition also holds. □

Theorem 1

Proof. Recall that $\Pi(\mu, s^*(\mu)) = t_r^*(\mu) + \frac{\partial t_r(\mu, s^*(\mu))}{\partial \mu}(\mu_0 - \mu)$. Taking derivative w.r.t μ, we have,

$$\frac{d\Pi(\mu, s^*(\mu))}{d\mu} = \frac{dt_r^*}{d\mu} - \frac{\partial t_r^*}{\partial \mu} + (\mu_0 - \mu)\frac{\partial^2 t_r^*}{\partial \mu^2} = \frac{\partial t_r^*}{\partial s}\frac{ds^*}{d\mu} + (\mu_0 - \mu)\frac{\partial^2 t_r^*}{\partial \mu^2} = \left[\frac{\frac{\partial t_r^*}{\partial s}\frac{ds^*}{d\mu}}{\frac{\partial^2 t_r^*}{\partial \mu^2}} + \mu_0 - \mu \right] \frac{\partial^2 t_r^*}{\partial \mu^2}$$
$$= \left[-\frac{(1-\mu)}{v_h}\mathbb{E}[v|q(s^*(\mu))] + \mu_0 - \mu \right] \frac{\partial^2 t_r^*}{\partial \mu^2}$$

Since t_r^* is concave on the domain $[\underline{\mu}^*, 0.5]$,[43] $\frac{\partial^2 t_r^*}{\partial \mu^2} < 0$. Let $\phi(\mu) = \frac{(1-\mu)}{v_h}\mathbb{E}[v|q(s^*(\mu))]$. The monotonicity of $\Pi(\mu, s^*(\mu))$ can be pinned down by the sign of $\mu_0 - \mu - \phi(\mu)$. In particular, if $\mu_0 - \mu > \phi(\mu)$, $\Pi(\mu, s^*(\mu))$ is decreasing in μ, otherwise, it's increasing in μ.

Claim 3. *$\phi(\mu)$ is decreasing and convex on $[\underline{\mu}^*, 0.5]$ and $\phi'(0.5) > -1$.*

We first prove this claim. Denote $w(\mu) := \mathbb{E}[v|q(s^*(\mu))]$. Then $\phi(\mu) = \frac{1-\mu}{v_h} w(\mu)$. It is easy to see that $w(\mu)$ is decreasing, since $q(s)$ is decreasing in s and $s^*(\mu)$ is increasing in μ. Besides, we can verify that $s^*(\mu)$ is concave,[44] hence $w(\mu)$ is convex. Note that

$$\begin{aligned}\phi'(\mu) &= -\frac{1}{v_h}[w(\mu) - (1-\mu)w'(\mu)] \\ &= -\frac{1}{v_h}\left[w(\underline{\mu}^*) + \int_{\underline{\mu}^*}^{\mu} w'(\mu)d\mu - (1-\mu)w'(\mu) \right]\end{aligned}$$

Since $w' < 0$ and $w'' > 0$, then $\int_{\underline{\mu}^*}^{\mu} w'(\mu)d\mu - (1-\mu)w'(\mu)$ is decreasing in μ and therefore $\phi'(\mu)$ is increasing on μ. That is, $\phi(\mu)$ is convex.

---

42 $\frac{d\Upsilon}{d\mu} = \frac{k(1-2\mu)\mu v_h + 2k(1-\mu)^2 v_l}{\lambda(\mu-1)\mu^3} < 0$ when $\mu \leq 0.5$.

43 $\frac{\partial^2 t_r^*}{\partial \mu^2} = \frac{k[(2\mu-1)\mathbb{E}(v|\mu) - (1-\mu)v_l]}{\lambda(1-\mu)^2\mu^3(v_h - v_l)} < 0$.

44 We can verify that $\frac{d^2 s^*}{d\mu^2}$ is proportional to $q(s^*(\mu))^2 M + \mu^2 N$, where $M \equiv (\mu(v_h - 4v_l) - 2\mu^2(v_h - v_l) + 2v_l)^2$ and $N \equiv (-2 + (5-4\mu)\mu)\mu v_h^2 + 2(1-\mu)^2(-3+2\mu)v_l v_h$. We can verify that $M > 0$, $N < 0$ and $M + N < 0$. Meanwhile $q(s^*(\mu)) < \mu$. Therefore $\frac{d^2 s^*}{d\mu^2} < 0$.

Denote $q^*(\mu) := q(s^*(\mu))$. Simplify $\phi'(0.5)$.

$$\phi'(0.5) = -\left[\frac{4(v_h - v_l)v_l q^*(0.5)^2}{v_h^2}(1 - q^*(0.5)) + \frac{1}{v_h}\mathbb{E}[v|q^*(0.5)]\right]$$

We can show that $\phi'(0.5)$ is decreasing in v_l. Hence plugging $v_l = 0$ and $v_l = v_h$ into $\phi'(0.5)$, we have $\phi'(0.5)|_{v_l=0} = -q^*(0.5) > -1$ and $\phi'(0.5)|_{v_l \to v_h} = -1$.[45] *Q.E.D*

Now we start to prove Theorem 1. Recall from Lemma 3, only for $s \in [q^{-1}(\underline{\mu}^*), Q^{-1}(0.5)]$, Stochastic Return is locally optimal. Hence, if we incorporate $q^{-1}(\mu_0) \le s \le Q^{-1}(\mu_0)$ and $\mu \le \mu_0$, then Stochastic Return could be an optimal solution only if

$$[q^{-1}(\underline{\mu}^*), Q^{-1}(0.5)] \cap [q^{-1}(\mu_0), Q^{-1}(\mu_0)] \neq \emptyset \text{ and } \mu_0 \ge \underline{\mu}^*$$

$$\iff \mu_0 \in [\underline{\mu}^*, Q(q^{-1}(\underline{\mu}^*))]$$

This can be easily observed from the graph below. In particular, when $\mu_0 \in [\underline{\mu}^*, 0.5)$, along the path of $t_r^*(\mu)$, the optimal stopping belief is constrained by $\mu \in [\underline{\mu}^*, \mu_0]$; while if $\mu_0 \in [0.5, Q(q^{-1}(\underline{\mu}^*))]$, the optimal stopping belief is constrained by $\mu \in [\underline{\mu}^*, \mu^*(Q^{-1}(\mu_0))]$. We distinguish two cases. First, $\phi'(\underline{\mu}^*) \ge -1$ and this implies that $\Pi(\mu, s^*(\mu))$ is quasi-convex. Second, $\phi'(\underline{\mu}^*) < -1$ and there exists a local maximum of $\Pi(\mu, s^*(\mu))$, which we can verify to be strictly worse than the revenue from Learning Deterrence. We establish the proof case by case.

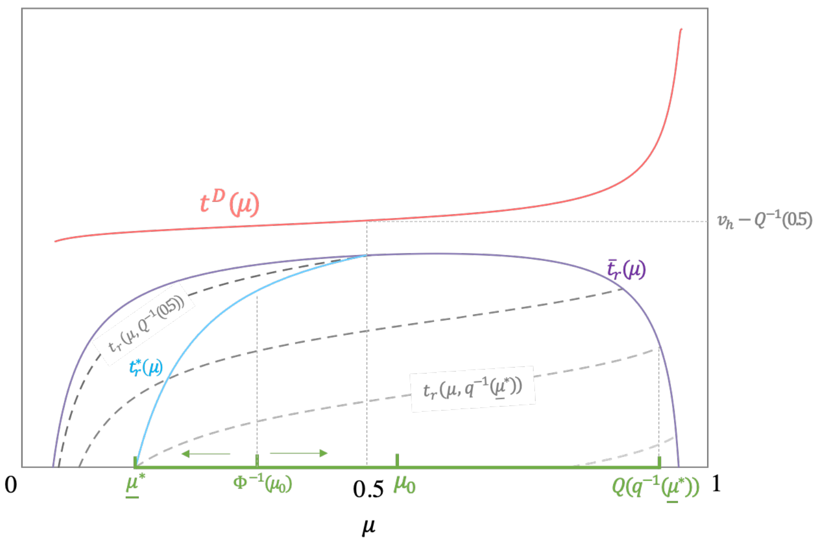

Figure 10: Expected revenue is quasi-convex

[45]Taking implicit differentiation w.r.t v_l for $\Pi_1(\mu, s^*(\mu))|_{\mu=0.5} = 0$, we have $\frac{ds^*(0.5)}{dv_l} = q^*(0.5) - 1 < 0$. Then $\frac{dq^*(0.5)}{dv_l} > 0$. Besides, $q^*(0.5) < 0.5$, then $\frac{d\phi'(0.5)}{dv_l} = \frac{1}{v_h^2}[(q^*(0.5)-1)(v_h + 4(v_h - 2v_l)q^*(0.5)^2) - (v_h - v_l)[v_h + 4v_l(2 - 3q^*(0.5))q^*(0.5)]\frac{dq^*(0.5)}{dv_l}] < 0$.

Case one: $\phi'(\underline{\mu}^*) \geq -1$. In most scenarios, this is true. Denote $\Phi(\mu) = \mu + \phi(\mu)$. Therefore when $\mu_0 \in [\Phi(\underline{\mu}^*), \Phi(0.5)]$, $\mu_0 - \mu$ single-crosses $\phi(\mu)$ from above, which is depicted in Figure 11. The black lines are the contour lines of $\mu_0 - \mu$ for different μ_0.

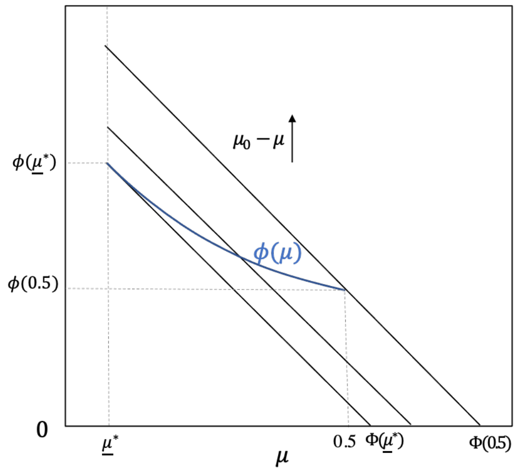

Figure 11: Expected revenue is quasi-convex

- If $\mu_0 < \Phi(\underline{\mu}^*)$, $\Pi(\mu, s^*(\mu))$ is increasing in μ, therefore Learning Deterrence is optimal. To verify it, when $\mu_0 < 0.5$, optimal stopping belief is μ_0 and inducing no learning via Stochastic Return is strictly dominated by Learning Deterrence. When $\mu_0 \geq 0.5$, we need to consider the feasible stopping belief $\mu \in [\underline{\mu}^*, \mu^*(Q^{-1}(\mu_0))]$ where $\mu^*(Q^{-1}(\mu_0)) < 0.5 < \mu_0$. However, we can show that Learning Deterrence is better than Stochastic Return that induces stopping at $\mu^*(Q^{-1}(\mu_0))$. To see this, suppose we ignore the constraint that $\mu \in [\underline{\mu}^*, \mu^*(Q^{-1}(\mu_0))]$, then the seller obtains larger revenue by inducing stopping at 0.5 with revenue as a weighted average between the selling price $v_h - Q^{-1}(0.5)$ and the return transfer $t_r(0.5, Q^{-1}(0.5))$. While with Learning Deterrence, she obtains revenue $t^D(\mu_0) > t^D(0.5) = v_h - Q^{-1}(0.5)$. Done.[46]

- If $\mu_0 \in [\Phi(\underline{\mu}^*), \Phi(0.5))$, $\Pi(\mu, s^*(\mu))$ is quasi-convex in μ. When $\mu_0 < 0.5$, optimal stopping belief is either $\underline{\mu}^*$ or μ_0, which implies the optimality between Free Return and Learning Deterrence. When $\mu_0 \geq 0.5$, we can still obtain the optimality between Free Return and Learning Deterrence by applying the same reasoning as above.

- If $\mu_0 \geq \Phi(0.5)$, $\Pi(\mu, s^*(\mu))$ is decreasing in μ. Hence Free Return is optimal.

[46]The magnitude between $\Phi(\underline{\mu}^*)$ and 0.5 is ambiguous. But it does not affect the above argument.

Case two: When $\phi'(\underline{\mu}^*) < -1$, there exists a local maximizer of $\Pi(\mu, s^*(\mu))$. Denote $r = \{\mu \in [\underline{\mu}^*, 0.5] : \phi'(\mu) = -1\}$. If $\mu_0 \in [\Phi(r), \Phi(\underline{\mu}^*)]$, there exists a unique local maximizer $r_1(\mu_0) = \{\mu \in [\underline{\mu}^*, r] : \phi(\mu) = \mu_0 - \mu\}$ (see Figure 12 for visualization). If $\mu_0 \notin [\Phi(r), \Phi(\underline{\mu}^*)]$, then the expected revenue is quasi-convex and the argument in case one validates. We want to show

$$\Pi(r_1(\mu_0), s^*(r_1(\mu_0))) < t^D(\mu_0), \text{ if } \mu_0 \in [\Phi(r), \Phi(\underline{\mu}^*)].$$

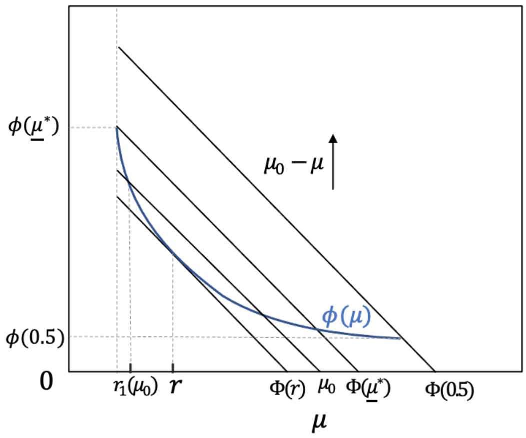

Figure 12: Expected revenue is not quasi-convex

To abuse the notation a little bit, we write $\Pi(\mu, s^*(\mu)|\mu_0)$ instead of $\Pi(\mu, s^*(\mu))$. Note that

$$\Pi(r_1(\mu_0), s^*(r_1(\mu_0))|\mu_0) < \Pi(r_1(\mu_0), s^*(r_1(\mu_0))|\Phi(\underline{\mu}^*)) < \Pi(\underline{\mu}^*, s^*(\underline{\mu}^*)|\Phi(\underline{\mu}^*)).$$

The first inequality comes from that Π is increasing in μ_0. The second inequality is because $\underline{\mu}^* = r_1(\Phi(\underline{\mu}^*))$, which is the maximizer of Π when $\mu_0 = \Phi(\underline{\mu}^*)$. Notice that,[47]

$$\Pi(\underline{\mu}^*, s^*(\underline{\mu}^*)|\Phi(\underline{\mu}^*)) = 0 + (\Phi(\underline{\mu}^*) - \underline{\mu}^*)\frac{\partial t_r(\mu, s^*(\mu))}{\partial \mu}|_{\mu=\underline{\mu}^*} = \mathbb{E}(v|\frac{k}{\lambda v_h})$$

It is obvious that $s = v_h - t_b < v_h$. Then $\mathbb{E}(v|\frac{k}{\lambda v_h}) < \mathbb{E}(v|\underline{\mu}) = t^D(\underline{\mu}) < t^D(\mu_0)$. The equality and the second inequality come from Proposition 2. We are done. □

Lemma 4

Proof. We can verify that $\Pi(q(s), s)$ is concave.

$$\frac{d^2\Pi}{ds^2} = \frac{2k\lambda(\mu_0 - 1)(k - \lambda v_h)}{(k - \lambda s)^3} < 0$$

[47] Recall $\underline{\mu}^* = \frac{k}{\lambda v_h} + (\frac{k}{\lambda v_h}(\frac{k}{\lambda v_h} + \frac{v_l}{v_h - v_l}))^{\frac{1}{2}}$.

Simply solve the optimization, we obtain:

$$s^F(\mu_0) = \frac{k}{\lambda} + \frac{\sqrt{k(\mu_0 - 1)\mu_0(k - \lambda v_h)}}{\lambda \mu_0},$$

$$\Pi^F(\mu_0) = \frac{-2\sqrt{k(\mu_0 - 1)\mu_0(k - \lambda v_h)} + k - 2k\mu_0 + \lambda\mu_0 v_h}{\lambda}.$$

Take derivative w.r.t μ_0.

$$\frac{ds^F}{d\mu_0} = \frac{k(k - \lambda v_h)}{2\mu_0\sqrt{k\lambda^2(\mu_0 - 1)\mu_0(k - \lambda v_h)}} < 0$$

Hence, the selling price $v_h - s^F(\mu_0)$ is increasing in μ_0. Furthermore, $\Pi^F(\mu_0)$ is increasing in μ_0 due to envelope theorem. □

Theorem 2

Proof. First, we need to verify that if $\Pi^F(\mu_0) \geq t^D(\mu_0)$, $q^{-1}(\mu_0) \leq s^F(\mu_0) \leq Q^{-1}(\mu_0)$. It obvious that if $\Pi^F(\mu_0) \geq t^D(\mu_0)$, $v_h - s^F(\mu_0) > t^D(\mu_0) = v_h - Q^{-1}(\mu_0)$ since the expected probability of successful sale is less than one with Free Return. Hence $s^F(\mu_0) \leq Q^{-1}(\mu_0)$ holds trivially. Moreover, if $\Pi^F(\mu_0) \geq t^D(\mu_0) > 0$, the expected probability of successful sale is positive, then $q(s^F(\mu_0)) < \mu_0$. Since $q(s)$ is decreasing, then $s^F(\mu_0) > q^{-1}(\mu_0)$. □

Proposition 3

Proof. Recall that $\Pi(q(s), s) = \frac{\mu_0 - \gamma/s}{1-\gamma/s}(v_h - s)$. By envelope theorem, we have:

$$\frac{d\Pi^F}{d\gamma} = \frac{v_h - s^F}{(s^F - \gamma)^2}(\mu_0 - 1)s^F < 0.$$

Hence Π^F is decreasing in γ.

To show that $t^D(\mu_0) = v_h - D(\mu_0)$ is increasing in γ, we want to show $D(\mu_0)$ is decreasing in γ. Recall that $\tilde{\mu}(\mu) = q(D(\mu))$. Take derivative w.r.t γ for both sides of $\mathbb{E}(v|\mu_0) - (v_h - D(\mu_0)) = V(\mu_0, D(\mu_0))$, we obtain:

$$\frac{1 - \mu_0}{1 - \tilde{\mu}(\mu_0)}\frac{dD}{d\gamma} = \frac{1 - \mu_0}{1 - \tilde{\mu}(\mu_0)} - 1 - (1 - \mu_0)\log\left[\frac{\mu_0/1 - \mu_0}{\tilde{\mu}(\mu_0)/1 - \tilde{\mu}(\mu_0)}\right] < 0.$$

Given that F is either empty or a close interval, it is immediate that if $\gamma_1 < \gamma_2$, then $F(\gamma_2) \subseteq F(\gamma_1)$. Note that $\underline{\mu}$ is the smaller root for $\mathbb{E}(v|\mu) - (v_h - q^{-1}(\mu)) = 0$. By implicit differentiation,

$$\frac{d\underline{\mu}}{d\gamma}(\frac{\gamma}{1 - \underline{\mu}} - \frac{\gamma}{\underline{\mu}}) = -1.$$

Hence $\frac{d\mu}{d\gamma} > 0$. Meanwhile, $\overline{\mu} = 1 - \underline{\mu}$, then $[\underline{\mu}(\gamma_2), \overline{\mu}(\gamma_2)] \subset [\underline{\mu}(\gamma_1), \overline{\mu}(\gamma_1)]$. $\square$

Proposition 4

Proof. First we calculate the limit of $t^D(\mu_0)$ when $\gamma \to 0$. Plugging $D(\mu_0) = \frac{\gamma}{\tilde{\mu}(\mu_0)}$ into equation (10) and multiply by γ.

$$-D(\mu_0) + \gamma \log \frac{\gamma}{D(\mu_0) - \gamma} = \frac{\gamma}{1 - \mu_0} + \gamma \log \frac{\mu_0}{1 - \mu_0} - (v_h - v_l)$$

When $\gamma \to 0$ and μ_0 does not converge to 0 or 1, the above equation converges to $v_h - D(\mu_0) = v_l$.[48] Hence $\lim_{\gamma \to 0} t^D(\mu_0) \to v_l$. For the expected revenue from Free Return,

$$\lim_{\gamma \to 0} \Pi^F(\mu_0) = \mu_0 v_h + \gamma(1 - 2\mu_0) - 2\sqrt{\gamma(1 - \mu_0)\mu_0(v_h - \gamma)} \to \mu_0 v_h$$

Therefore when $\gamma \to 0$, the seller is indifferent between Learning Deterrence and Free Return at $\mu_0 = \frac{v_l}{v_h}$. Since the above limit of $t^D(\mu_0)$ may fail when $\mu_0 \to 0$ or $\mu_0 \to 1$, we have to verify the extreme prior case that $\lim_{\gamma \to 0} [\underline{\mu}, \overline{\mu}] \to [0, 1]$. Plugging $\underline{\mu} = \frac{1}{2}\left(1 - \sqrt{1 - 4\gamma/(v_h - v_l)}\right)$, we have

$$\lim_{\gamma \to 0} \gamma \log \frac{\underline{\mu}}{1 - \underline{\mu}} = \gamma \log \frac{1 - \sqrt{1 - 4\gamma/(v_h - v_l)}}{1 + \sqrt{1 - 4\gamma/(v_h - v_l)}} \to 0.$$

Hence $\lim_{\gamma \to 0} t^D(\underline{\mu}) \to v_l$. Thus, when $\mu_0 < \frac{v_l}{v_h}$, the seller's expected revenue from the optimal mechanism converges to v_l.

Plugging $\overline{\mu} = \frac{1}{2}\left(1 + \sqrt{1 - 4\gamma/(v_h - v_l)}\right)$, we have

$$\lim_{\gamma \to 0} \gamma \log \frac{\overline{\mu}}{1 - \overline{\mu}} = \gamma \log \frac{1 + \sqrt{1 - 4\gamma/(v_h - v_l)}}{1 - \sqrt{1 - 4\gamma/(v_h - v_l)}} \to 0$$

$$\lim_{\gamma \to 0} \frac{\gamma}{1 - \overline{\mu}} = \frac{\gamma}{1 - \sqrt{1 - 4\gamma/(v_h - v_l)}} \to v_h - v_l$$

Hence $\lim_{\gamma \to 0} \overline{\mu} \to 1$, $\lim_{\gamma \to 0} t^D(\overline{\mu}) \to v_h$, and $\lim_{\gamma \to 0} \Pi^F(\overline{\mu}) = v_h$. If $\mu_0 \ll 1$, $\lim_{\gamma \to 0} \Pi^F(\mu_0) > \lim_{\gamma \to 0} t^D(\mu_0)$. Then when $\mu_0 \geq \frac{v_l}{v_h}$, the seller's expected revenue converges to $\mu_0 v_h$. $\square$

[48] $\lim_{\gamma \to 0} \gamma \log \frac{\gamma}{D(\mu_0) - \gamma} = 0$