Optimal Refund Mechanism*

Qianjun Lyu†

February 4, 2022

Abstract

This paper studies the optimal refund mechanism when an uninformed buyer can privately acquire information about his valuation over time. In principle, a refund mechanism can specify the odds that the seller requires the product returned while issuing a (partial) refund, which we call stochastic return. It guarantees the seller a strictly positive minimum revenue and facilitates intermediate buyer learning. In the benchmark model, stochastic return is always sub-optimal. The optimal refund mechanism takes simple forms: the seller either deters learning via a well-designed non-refundable price or encourages full learning and escalates price discrimination via free return. This result is robust to both positive learning and negative learning framework.

Key words: buyer learning, refund contract, sequential information acquisition, information design, screening.

*This paper was previously circulated under the title, "Learning Deterrence vs. Learning Encouragement: Optimal Pricing and Return Policy."

†Faculty of Business and Economics, University of Hong Kong. Email: lyu2014@hku.hk. I would like to thank my advisor Wing Suen for insightful and continual guidance. For valuable suggestions, I would also like to thank Pak Hung Au, Bard Harstad, Wei He, Jin Li, Yunan Li, Xianwen Shi, Balázs Szentes, Jianrong Tian, Dong Wei, Yimeng Zhang and Chen Zhao. All errors are mine.

1 Introduction

The rise of the Internet clears the way for consumers to acquire product information. Even before purchase, there are lots of information available on the Internet and social media that can help the consumers to make better decisions. However, whether it is necessary to acquire information; if yes, how much information the consumers should acquire, clearly depend on the pricing and return policy. For example, if the seller does not allow a return, then the consumer tends to make a more cautious purchase as he will acquire all necessary information before purchase; conversely, if the seller offers a free return, then there will be no regret for uninformed purchase. In this sense, a refund mechanism determines the buyer's value from learning. From the sellers' perspective, she can indirectly control the buyer's endogenous learning by designing different refund mechanisms, which will eventually affect the buyer's learning outcomes and then affect the seller's expected sales revenue.

This paper studies the revenue-maximizing refund mechanism anticipating that the buyer privately acquires information about his true valuation over time. A refund mechanism specifies a product's price and its return policy. In general, the return policy could take many different formats. Free return (return with a full refund) and no return are commonly used in practice. Moreover, a seller can offer a partial refund while receiving a return request. For example, airlines usually charge a fixed fee for a ticket refund. More surprisingly, e-commerce retailers such as Amazon sometimes issue a refund without requiring a product return.

Given the flexibility in designing return policies, in principle, the seller can allow the buyer to *keep* the item with some probability while issuing a refund. This generates a positive trading surplus upon a return. Moreover, this guarantees the seller a strictly positive minimum revenue since the buyer is willing to accept a partial refund in exchange for positive odds to keep the item. We call such a return policy *stochastic return*. Assuming quasi-linear consumer preferences, we can represent a return policy as (1) the probability that the seller requires the product returned and (2) the (expected) refund paid back to the buyer. This characterization can capture all the above-mentioned return policies.

For concreteness, consider a seller (she) selling one unit of an indivisible good to a buyer (he). The buyer's valuation could be either high or low, $v_h > v_l \geq 0$, and we normalize the seller's opportunity cost to be zero so that the first-best solution requires immediate consumption without learning. The buyer is initially uninformed about his true product valuation, and we interpret this uncertainty in product valuation as coming from match-specific factors so that the seller is symmetrically uninformed ex-ante. Therefore, the seller's major concern is to design a mechanism to implement some ideal amount of buyer learning.

The buyer can privately acquire information both before and after purchase. For example, a graphic designer, intending to buy an iPad, is initially uncertain about whether it is good for drawing and image editing. He can explore related information online or visit the Apple store to experience the product. Or he can purchase the product first and evaluate it afterwards. We adopt positive Poisson learning in the main model. Specifically, by exerting costly effort, e.g., spending time acquiring information, good news arrives according to some Poisson rate if the true valuation is high, otherwise no news arrives if the true valuation is low. Moreover, the learning rate after purchase is weakly higher than the learning rate before purchase since the information attainable before purchase is still attainable after purchase. Nevertheless, with the spread of information on the Internet, the consumer can obtain more and more instructive information before purchase, rendering the extra information generated by personal experience after purchase smaller. As a benchmark, we focus on the case where the learning rate is the same before and after purchase.

At the outset, the seller commits to a refund mechanism, after which, the buyer decides how much information to acquire and makes his purchase and subsequent return decision based on the information outcomes. Essentially, a refund mechanism offers the buyer two options: either to consume the item and obtain the consumption utility, or to return it and obtain the return payoff. Therefore, it creates the buyer an option value, which eventually affects the buyer's learning outcomes.

Building our model under the exponential-bandit framework (Keller, Rady and Cripps (2005)) allows us to disentangle the impacts of price and return policy on the buyer's learning behavior. Specifically, the price affects the buyer's incentive to learn and thereby determines his expected trading surplus. The return policy determines the total amount of information that the buyer optimally acquires, since the buyer makes the stopping decision comparing the return payoff and the continuation value from learning.

To find the revenue-maximizing refund mechanism, we first characterize the complete set of buyer learning outcomes that are inducible under any arbitrary refund mechanism. Next, we find the corresponding optimal mechanism that implements each learning outcome. Last, we recast the seller's objective function as the maximization over the inducible buyer learning outcomes rather than the feasible refund mechanisms. It reduces the dimension of the seller's choice variable.

The set of inducible buyer learning outcomes can be segmented into three groups: full learning, partial learning, and no learning. *Full learning* refers to the scenarios where the buyer stops learning when there is *zero* continuation value from learning. It can be implemented by free return. *Partial learning* refers to the scenarios where the buyer stops learning when there is still *positive* continuation value from learning. To implement it, the seller has to offer the buyer positive odds to keep the item upon return to compensate the opportunity information rent that the buyer could have enjoyed if he continued to learn. Finally, *No learning* refers to the case that the buyer consumes the product immediately without learning. To achieve this, the seller does not allow a return, but she carefully designs a price making the buyer just indifferent between consuming the item and continuing to learn.

Our main result in the benchmark model suggests that inducing partial learning is always sub-optimal,[1] implying the optimality of a deterministic mechanism. That is, if the seller allows a return, she requires the buyer to return the product with probability one. Otherwise, she does not allow a return. This bang-bang solution comes from the result that the seller's revenue is quasi-convex in the buyer's stopping belief if the price is also optimally adjusted. Intuitively, whenever the seller wants to extend the buyer's learning process–induce a lower stopping belief–in order to increase the odds of a successful sale, she can further benefit by raising the price simultaneously. Thus, it reinforces her incentive to drive down the stopping belief. Conversely, when the seller tends to increase the buyer's stopping belief to guarantee a larger minimum revenue (the revenue she obtains after refund), she can further raise this minimum revenue by lowering the price, reinforcing her incentive to increase the stopping belief.

Hence, the revenue-maximizing mechanism either prevents the buyer from private

[1] Note that in our problem, the seller's objective function is not linear in the allocation probability upon return as the buyer's stopping belief endogenously depends on the allocation probability.

learning or encourages full learning via free return. The optimality between them depends on the buyer's prior belief, which measures how much the buyer values information ex-ante and how optimistic the buyer initially is. Specifically, if the buyer is well-informed ex-ante, i.e., his prior belief is close to 0 or 1, then information is barely valuable to him. Therefore, the seller can induce immediate consumption by decreasing the price just a little to capture a large fraction of the first-best allocation surplus. However, if the buyer's prior belief becomes more uncertain, then information values a lot. Therefore, the seller must decrease the price significantly to deter learning, which makes encouraging learning more appealing as it avoids the compensation for the buyer's opportunity information rent. In other words, the seller significantly raises the price to encourage learning while allowing a free return. However, encouraging learning causes inefficient allocation as the buyer would eventually return the product. Nevertheless, this event becomes rare when the prior is more optimistic. As a result, the seller optimally allows free return if the buyer's prior belief is less extreme but relatively optimistic. Otherwise, the optimal mechanism prevents learning.

Interestingly, though the buyer enjoys a larger information rent if his prior belief is less extreme, he can only benefit from it if the seller deters learning. In contrast, free return causes a severe decline in the buyer's trading surplus as the seller escalates price discrimination. It means that the buyer takes the cost of learning and inefficient allocation when the seller encourages him to learn.

When learning after purchase is more efficient than before purchase, the seller optimally charges a cancellation fee (equivalent to partial refund) to extract the extra information rent from post-purchase learning. Nevertheless, the main result is robust in the sense that deterministic mechanism is optimal if the difference between the before-purchase and post-purchase learning is not very large. We also discuss the scenarios where the learning rate after purchase converges to infinity so that the buyer can almost learn his true valuation immediately. In another extension, we discuss negative Poisson learning (no-news-is-good-news). The optimal refund mechanism turns out to have the same structure as in the benchmark model. However, with this learning technology, the seller can use free return to fully extract the buyer's surplus. In other words, the buyer receives *zero* surplus if the seller optimally allows free return.

Related literature. Our paper relates to the sequential screening literature. Courty and Li (2000) study the refund contract to price discriminate the buyer who has imperfect private information ex-ante but observes his true valuation after contracting. In contrast, we consider symmetric ex-ante information and study the refund contract to elicit the buyer's ex-post private information, which is the buyer's endogenous learning outcome. Krähmer and Strausz (2015) impose the ex-post participation constraints in the standard sequential screening model to capture the mandated consumer withdrawal right. In our paper, whether to offer the consumer an option of ex-post participation is an endogenous choice of the seller.

There is a growing literature on mechanism design incorporating the buyer's endogenous information acquisition. For example, Shi (2012) and Mensch (2020) study mechanism design when the buyer can privately acquire costly information. Shi (2012) adopts rotational-ordered information technology, and Mensch (2020) discusses flexible information acquisition, with cost as the expected difference in a posterior-separable measure of uncertainty. Mensch (2020) characterizes the set of implementable mechanisms to screen the buyer with different interim information. We adopt a similar method; however, our exponential bandit specification with additive time cost allows us to analyze how the seller's optimal mechanism varies with the buyer's initial belief, which cannot otherwise be accommodated in the flexible information cost framework.[2] In terms of sequential buyer learning,[3] Lang (2019) and Pease (2020) investigate the seller's pricing policy when the buyer can acquire information over time.

The closest work to our study are Matthews and Persico (2007), Board (2007) and Daley, Geelen and Green (2021), which analyse sequential mechanism with endogenous buyer learning. Specifically, Matthews and Persico (2007) discuss the seller's optimal choice of price and refund, anticipating that the buyer can acquire perfect information at a fixed cost before purchase. Therefore, stochastic mechanism does not have a bite since the buyer either acquires perfect information or no information. We differ by discussing imperfect learning so that the seller has much greater

[2]In our model, the cost of the same Blackwell experiment is the same for different prior beliefs, which is not true for flexible information. There does not exist a unified measure of uncertainty, regardless of the prior beliefs, that can represent the additive time cost of Poisson signals: see Appendix A of Mensch (2020) and Pomatto, Strack and Tamuz (2019).

[3]See Bonatti (2011), Bergemann and Valimaki (2000) and Bergemann and Valimaki (1996).

flexibility in manipulating the buyer's learning behavior. Board (2007) and Daley, Geelen and Green (2021) investigate option contracts where the winning bidder can choose whether to execute the option after collecting new information. In Board (2007), a winning bidder can choose whether to use an asset at a contingent fee or give up the upfront payoff and quit the market. Daley, Geelen and Green (2021) discuss due diligence in M&A, wherein after the acquirer agrees on the price with the target firm, he has the option not to execute the contract. Both papers focus on deterministic execution, while in contrast, we allow stochastic execution.

Our paper also relates to mechanism design with information discrimination. Li and Shi (2017) allow the seller to disclose different additional information to different types of buyers to enhance price discrimination. Wei and Green (2020) study the optimal information discrimination incorporating the buyer's ex-post participation. Instead of the seller restricting the buyer's learning process, Roesler and Szentes (2017) adopt a robustness perspective by allowing the buyer to acquire costless information anticipating its impact on the seller's pricing decision. Hinnosaar and Kawai (2020) investigate robust refund mechanism to capture the situations where the seller is unsure about the buyer's private information prior to purchase. Johnson and Myatt (2006) introduce rotations of demand curves to capture the dispersion of consumer valuations and discuss how seller profits change with the level of dispersion.

The literature offers several complementary economic rationales for the optimality of refund contracts. Che (1996) shows that the seller optimally insures risk-averse buyers by offering a generous refund. Inderst and Ottaviani (2013), Shieh (1996) and Inderst and Tirosh (2015) discuss the role of refund as a signaling device to guarantee credible sales talk, product quality, and personal fit.

The remainder of this paper is organized as follows. Section 2 describes the model. Section 3 characterizes the set of inducible buyer learning outcomes. Section 4 and 5 discuss the implementable mechanism and the seller's optimization problem. Section 6 analyses the optimal mechanism. Section 7 discusses more efficient post-purchase learning. Section 8 studies negative Poisson learning. Section 9 provides several minor extensions. Section 10 concludes.

2 Model

A seller (female) sells one unit of indivisible goods to a risk-neutral buyer (male). The buyer is initially uninformed about his true product valuation, either high or low, $v_h > v_l \geq 0$. The seller is symmetrically uninformed, with μ_0 being the common prior belief that the product valuation is high. We use μ to represent the buyer's posterior belief after learning and sometimes call this the buyer's type. A type-μ buyer's expected value of the product is $\mathbb{E}(v|\mu) := \mu v_h + (1-\mu)v_l$. Note that the buyer's type evolves depending on the learning process; we use τ to denote time and write $\mu(\tau)$ when needed. We focus on the scenario where efficiency requires trade with probability one, and therefore normalize the seller's opportunity cost to 0. There is no cost of production or return. We assume that neither party discounts over time.[4]

The seller commits to a refund mechanism, which specifies (1) a price $t_b \geq 0$, which is the transfer made from the buyer to the seller at the time of purchase; and (2) a return policy that describes the probability that the buyer is required to return the item and the (expected) refund paid back to the buyer. Given that the buyer is assumed to be risk-neutral, only the expected refund matters. For the sake of exposition, we use (x_r, t_r) to denote a return policy. Precisely, $x_r \in [0,1]$ is the probability that the buyer is allowed to *keep* the item after requesting a return. The reader can interpret x_r as the allocation rate at return. $t_r \in [0, t_b]$ is expected final payment made from buyer to seller if the buyer requests a return. We call it the return transfer later on. A typical refund mechanism is characterized by $\{t_b, (x_r, t_r)\}$. Under this mechanism, the buyer pays the price t_b at the time of purchase. If the buyer requests a return, then the seller applies to a public randomization device: with probability x_r, she allows the buyer to keep the item, with the remaining probability, she requires the buyer to return it. Meanwhile, the seller pays the (expected) refund $t_b - t_r$ regardless of whether the item eventually returns to her.

Given this notation, a No Return mechanism can be represented as $\{t_b, (1, t_b)\}$. In particular, $x_r = 1$ means the buyer cannot return the item and therefore always

[4] Suppose there is discounting and the seller allows a free return. Then the seller gains a positive payoff even if the buyer returns the item and receives a full refund. However, the time between purchase and return is usually not very long, so we assume no discounting to get rid of this issue.

has to keep it. Free Return can be represented as $\{t_b, (0,0)\}$, so that $x_r = 0$ as the buyer can return the product for a refund. Stochastic Return requires $x_r \in (0,1)$, so that the buyer can keep the item with strictly positive probability even upon obtaining a refund. We capitalize the first letter of a return policy to represent a refund mechanism and emphasize that the price can vary while fixing the return policy. Without loss of generality, we assume $v_h - t_b > v_h x_r - t_r$, i.e., a high-value buyer purchases the item without requesting a further return. The buyer's outside option is normalized to zero.

A type-μ buyer's payoff is realized when he consumes the item. If so, he cannot request a return, regardless of the return policy. In particular, a type-μ buyer obtains expected utility $\mathbb{E}(v|\mu) - t_b$ if he purchases the item without requesting a return, or $\mathbb{E}(v|\mu)x_r - t_r$ if he requests a return. Let $\mathbf{B}_\tau$ be the indicator function for whether a purchase has occurred up to and *including* time τ. Hence, the time of purchase is $\tau_b = \min\{\tau : \mathbf{B}_\tau = 1\}$. Analogously, $\mathbf{R}_\tau$ denotes the indicator function for whether a return has occurred up until time τ, and the time that the buyer requests a return is $\tau_r = \min\{\tau : \mathbf{R}_\tau = 1\}$. Naturally, $\tau_r \geq \tau_b$. The seller's revenue Π is expressed as follows:

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

The buyer can acquire information both before and after purchase. Specifically, we adopt the exponential bandit framework, and in the main model, we consider the case of “no-news-is-bad-news". If the buyer pays a fixed flow cost k to acquire information, then good news arrives according to some Poisson rate if his true valuation is v_h and no news arrives if his true valuation is v_l. We denote λ_B (λ_P) as the before-purchase (post-purchase) learning rate.

We assume $\lambda_P \geq \lambda_B$ since the information attainable before purchase is still attainable after purchase. However nowadays, with the spread of information on the Internet and social media, the consumer can obtain more and more instructive information before purchase, and the extra information generated by personal experience after purchase becomes smaller. Besides, many retailers, such as Apple store, allow the consumer to experience their products at the off-line store, therefore there is not a large difference between the information attainable to the consumer before and after purchase. Thus, in the benchmark model, we focus on the case where

$\lambda_P = \lambda_B = \lambda$. In section 7, we discuss the scenarios where $\lambda_P > \lambda_B$ and let $\lambda_P \to \infty$ to capture the case where the buyer can learn his true valuation immediately.

Given $\lambda_P = \lambda_B = \lambda$, the buyer's belief evolves according to the following law of motion if no Poisson jump occurs:

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

Otherwise, if good news arrives, his belief jumps to one. Without loss of generality, if the seller does not allow a return, the buyer acquire information before purchase; conversely, if the seller allows a return, we assume the buyer purchases the item first and acquires information afterwards.

3 A No Return Benchmark

In this section, we study the buyer's learning strategy if the seller does not allow a return. The result serves as a building block for the derivation of subsequent results. Because the buyer can always give up the option to return, such as treating all mechanisms as if no return is allowed. Therefore, the buyer's valuation from learning under a No Return mechanism imposes a lower bound on his expected surplus.

Denote $s := v_h - t_b$ as the net consumer's surplus upon the arrival of good news. It is endogenously chosen by the seller and it determines the value of experimentation, $V^0(\mu(\tau), s)$, which is characterized by the Bellman equation:

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

At time τ, the buyer can walk away or purchase the item. If he continues to learn for an interval of time $d\tau$ then, with probability $\mu(\tau)\lambda d\tau$, good news arrives, and he purchases the item; with the remaining probability, no news arrives, and his belief decreases to $\mu(\tau + d\tau)$. Conditional on learning, the Bellman equation leads to this differential equation:

$$(1 - \mu)\mu\lambda V_1(\mu, s) + \mu\lambda V(\mu, s) = \mu\lambda s - k, \tag{ODE}$$

where $V_1(\mu, s)$ denotes the partial derivative with respect to the first argument. Conventionally, for a fixed s, there exists two cutoff beliefs: the quitting belief $q(s)$

and the trial belief $Q(s)$, with $q(s) \leq Q(s)$, which determine the buyer's optimal learning strategy. That is, he continues to learn when his belief falls between the two cutoffs; otherwise, he does not learn. The quitting belief $q(s)$ is determined by the standard value matching and smooth pasting conditions,[5] and it adopts a closed-form solution:

$$q(s) = \frac{k}{\lambda s}.$$

The trial belief is the value of belief above which the buyer strictly prefers immediate consumption to acquiring information:

$$Q(s) = \{\mu : V(\mu, s) = \mathbb{E}(v|\mu) - (v_h - s)\}. \tag{3}$$

With slight abuse of notation, in equation (3) and henceforth, we use $V(\mu, s)$ to denote the solution of (ODE) with boundary point $(q(s), 0)$.

The construction above involves one implicit assumption: when the buyer stops learning at $q(s)$, he prefers to quit the market than accept the price. Specifically,

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

We call it the Learning-Feasibility constraint. If it fails, no learning can be induced because learning has no value when it does not affect the purchase decision. If this constraint fails for all $s \in [0, v_h - v_l]$,[6] the buyer then considers learning sub-optimal regardless of the price. The seller can set a No Return mechanism with a price equal to the ex-ante expected value of the product and capture the entire allocation surplus $\mathbb{E}(v|\mu_0)$. To avoid this trivial result, throughout the paper, we assume that there exist two distinct roots $\underline{s} < \overline{s}$ that the Learning-Feasibility constraint binds, which is equivalent to the assumption below.

Assumption: $(v_h - v_l)\lambda > 4k$.

That is, if the learning cost is not very high or the Poisson rate is not very low, the buyer would consider learning valuable for some values of s.

Proposition 1. *If $s \notin [\underline{s}, \overline{s}]$, $V^0(\mu, s) = \max\{0, \mathbb{E}(v|\mu) - (v_h - s)\}$; and if $s \in [\underline{s}, \overline{s}]$,*

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

[5] $q(s) = \{\mu : V_1(\mu, s) = 0 \text{ and } V(\mu, s) = 0\}$.

[6] $s \notin [0, v_h - v_l]$ means that the price is in between v_l and v_h.

Proposition 1 characterizes the buyer's learning strategy under no return while varying the surplus s through the price. For sufficiently high or sufficiently low surplus s, learning is sub-optimal regardless of the buyer's belief. With moderate $s \in [\underline{s}, \overline{s}]$, the standard results of exponential bandit apply: When the buyer's prior belief falls into $[q(s), Q(s))$, he optimally learns until either good news arrives and he purchases the item, or no news arrives for a sufficient amount of time and he walks away at the quitting belief $q(s)$.

Nevertheless, the above characterization of the buyer's learning strategy only applies to No Return mechanisms. If the seller varies the return policy, the buyer changes his learning strategy accordingly. Therefore, Lemma 1 is an essential simplification result as it pins down the buyer's value function under any optimal refund mechanism.

Lemma 1. *For a fixed s, all optimal refund mechanisms provide the buyer with a expected trading surplus of $V^0(\mu_0, s)$.*

That is, under any optimal refund mechanism, the buyer obtains the same continuation value as if the mechanism prohibited a return. To see this, for a fixed price, if the seller designs a benevolent return policy that provides the buyer with a continuation value strictly larger than $V^0(\mu_0, s)$, she can then increase the return transfer t_r and adjust the return allocation rate x_r properly without affecting the total amount of information that the buyer acquires, which implies a profitable deviation.

Given this simplification result, we are able to characterize the set of buyer learning outcomes that are inducible under any potentially optimal refund mechanism. Figure 1 plots the quitting belief $q(s)$ and trial belief $Q(s)$ against s. These two beliefs are decreasing in s and coincide at the two boundaries.[7] Denote $\underline{\mu} = q(\bar{s}) = Q(\bar{s})$ and $\overline{\mu} = q(\underline{s}) = Q(\underline{s})$.

[7] The quitting belief is decreasing in s because the buyer optimally learn for a longer time if the benefit from good news becomes larger. The trial belief is also decreasing in s. Because if the seller increases s by one unit, then the consumption utility increases by one unit, but the increment of the buyer's continuation value is smaller than one unit. $q(s) = Q(s)$ at the boundaries is implied by the Learning-Feasibility constraint and the definitions of $q(s)$ and $Q(s)$.

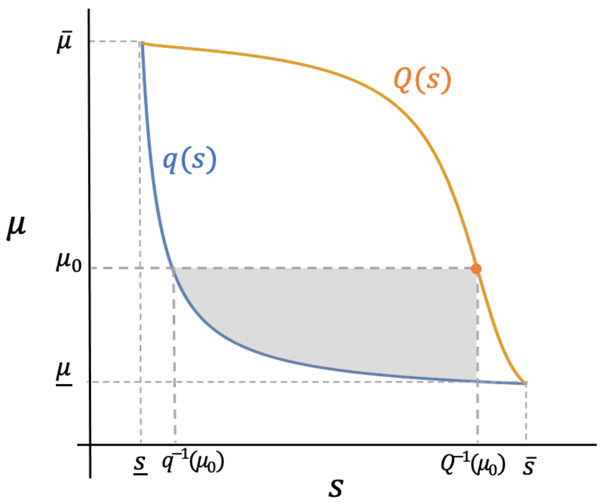


Figure 1: Inducible learning outcomes

The shaded area in Figure 1 is the set of inducible learning outcomes for a buyer with prior belief μ_0. Specifically, an *inducible learning outcome* is a pair of (s, μ) such that for a given s, there exists a return policy that induces the buyer to stop learning and request a return at posterior belief $\mu \in [q(s), Q(s)]$. Therefore, we require $s \in [q^{-1}(\mu_0), Q^{-1}(\mu_0)]$ so that the type-μ_0 buyer prefers to learn at the prior belief. Meanwhile, the stopping belief $\mu \leq \mu_0$ since no news is bad news.

Before we move on, the corollary below describes an extreme case, $\mu_0 \notin [\underline{\mu}, \overline{\mu}]$, such that the buyer is sufficiently informed upfront and deems learning to be sub-optimal. Then the seller can extract the entire allocation surplus by setting a non-refundable price equal to the ex-ante expected valuation of the good, which leaves the buyer zero trading surplus. No learning is induced on path.

Corollary 1. *If $\mu_0 \notin [\underline{\mu}, \overline{\mu}]$, the optimal mechanism is No Return, with $t_b = \mathbb{E}(v|\mu_0)$ and $(x_r, t_r) = (1, t_b)$.*

However, if the buyer has a less extreme prior belief, $\mu_0 \in [\underline{\mu}, \overline{\mu}]$, learning becomes a valuable option to him, which prevents the seller from capturing the entire allocation surplus. Then the seller faces a non-trivial tension between deterring learning and encouraging learning. In Section 4 and 5, we find the refund mechanism that can implement each inducible learning outcome in the shaded region, and then let the seller maximize her expected revenue over the inducible learning outcomes.

4 Learning Deterrence and Free Return

In this section, we characterize the refund mechanisms that can implement the boundaries of the shaded area. In particular, the No Return mechanism named *Learning Deterrence* implements the intersection of the shaded area and the orange curve $Q(s)$ (the orange dot in Figure 1), and Free Return mechanisms implement the intersection of the shaded area and the blue curve $q(s)$.

Learning Deterrence is a No Return mechanism with price $t^D := v_h - Q^{-1}(\mu_0)$.[8] Under this mechanism, the type-μ_0 buyer is just indifferent between acquiring information and consuming the item immediately (see the orange dot in Figure 1). We let the buyer break indifference by purchasing the item immediately so that to achieve efficient allocation. Notice that to deter learning, the seller has to lower the price so as to give away part of the allocation surplus to compensate the buyer until the value of information becomes non-positive.[9] In other words, the buyer's expected trading surplus, which is just the consumption utility in this case, equals his continuation value from learning, i.e., $\mathbb{E}(v|\mu_0) - t^D(\mu_0) = V(\mu_0, Q^{-1}(\mu_0))$. Furthermore, the joint surplus equals the full allocation surplus $\mathbb{E}(v|\mu_0)$, and the seller obtains revenue $\Pi^D(\mu_0) = t^D(\mu_0)$.

A **Free Return** mechanism $\{t_b, (0, 0)\}$, with $s \in [q^{-1}(\mu_0), Q^{-1}(\mu_0)]$, encourages the buyer to acquire information. Specifically, the type-μ_0 buyer continues to learn until good news arrives or no news arrives and his posterior belief falls to $q(s)$. By varying s, the seller can induce different quitting beliefs and thereby induce different amounts of buyer learning. In Figure 1, the intersection of the shaded region and the blue curve $q(s)$ represents the learning outcomes that can be induced by Free Return mechanisms. A common property of Free Return is that the buyer stops learning when the continuation value from learning is 0. We define it as *full learning* since it is the largest amount of information acquisition that the seller can induce when the price is fixed.

Given the flexibility of varying price while allowing a free return, the seller faces a

[8] Note that any non-refundable prices strictly lower than t^D can induce immediate consumption, but the seller then has an incentive to increase the price.

[9] The value of information refers to the difference between the value function and the payoff from optimally choosing between purchasing and walking away.

trade-off between increasing the price and reducing the return rate. The revenue-maximizing Free Return mechanism is determined by the constrained optimization problem ($\mathcal{F}$) below:

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

Note that $\frac{\mu_0 - q(s)}{1-q(s)}$ is the ex-ante probability that good news arrives before the buyer's belief falls below $q(s)$. Optimization over Free Return mechanisms is mechanical. The unconstrained optimization admits a closed-form solution. We denote the unconstrained maximizer as $s^F(\mu_0)$ and the corresponding revenue as $\Pi^F(\mu_0)$.

Remark: (Robustness) Conditional on Lemma 1, if the seller sets $s = Q^{-1}(\mu_0)$, type-μ_0 buyer is indifferent between learning and immediate consumption regardless of the return policy, as long as we assume the seller-preferred tie-breaking rule. Nevertheless, We choose Learning Deterrence (no return) to implement immediate consumption for the sake of robustness. Specifically, if the post-purchase learning rate is just slightly higher than the before-purchase learning rate, then if the seller allows a return while fixing $s = Q^{-1}(\mu_0)$, the buyer strictly prefers to purchase the item and acquire information afterward. It causes inefficient allocation and strictly reduces the seller's expected revenue. For a similar reasoning, we choose Free Return to encourage full leaning.

Let $\Pi^*(\mu_0)$ be the expected revenue from an optimal refund mechanism. Given that both Learning Deterrence and Free Return are feasible mechanisms, if $\mu_0 \in [\underline{\mu}, \overline{\mu}]$,

$$\Pi^*(\mu_0) \geq \max\{\Pi^D(\mu_0), \Pi^F(\mu_0)\}.$$[10]

We close this section by discussing the welfare properties of Learning Deterrence and the revenue-maximizing Free Return when the prior varies.

Proposition 2. *Under Learning Deterrence and $\mu_0 \in [\underline{\mu}, \overline{\mu}]$, the buyer's trading surplus $V(\mu_0, Q^{-1}(\mu_0))$ is non monotone and single-peaked in μ_0, while the seller's revenue $\Pi^D(\mu_0)$ is increasing in μ_0. The joint surplus is $\mathbb{E}(v|\mu_0)$.*

$$\textit{If } \mu_0 = \underline{\mu}, \overline{\mu}, \quad \textit{then} \quad V(\mu_0, Q^{-1}(\mu_0)) = \mathbb{E}(v|\mu_0) - \Pi^D(\mu_0) = 0.$$

[10] If $\Pi^{\mathcal{F}}(\mu_0) \geq \Pi^D(\mu_0)$, then we can verify that $\Pi^{\mathcal{F}}(\mu_0)$ adopts the same expression as $\Pi^F(\mu_0)$.

Learning Deterrence is different from the extreme case stated in Corollary 1, as the buyer must be induced to give up his option to learn, and this option is valuable when he is not very-well informed ex-ante, i.e., $\mu_0 \in [\underline{\mu}, \overline{\mu}]$. Therefore, to prevent the buyer from private learning, the seller has to sufficiently lower the price so that accepting the price is more attractive for the buyer than acquiring information. When the prior belief moves to more intermediate region, the buyer enjoys larger benefits from learning and thereby the seller must give away a larger amount of allocation surplus for the buyer's compensation if she wants to deter learning. This hints at the non-monotonicity of the buyer's trading surplus.

Proposition 3. *Under Free Return and $\mu_0 \in [\underline{\mu}, \overline{\mu}]$, both the price $v_h - s^F(\mu_0)$ and revenue $\Pi^F(\mu_0)$ are increasing in μ_0.*

Under Free Return, a more optimistic prior belief increases the probability of a successful sale, and the seller optimally puts more weight on getting a higher price and less weight on raising the probability of making a sale, rendering both the price and the revenue increasing in μ_0.

5 Partial Learning

In this section, we study the interior region of the inducible learning outcomes. That is, instead of encouraging the buyer to perform full learning or prevent the buyer from private learning, the seller can induce the buyer to stop at any intermediate belief in the shaded region in Figure 1. We define *partial learning* as the buyer stops learning when there is still positive continuation value from learning. To achieve this, the seller must provide the buyer with positive allocation rate at return in order to compensate the buyer's opportunity information rent which he could have been enjoyed if he continued to learn. Therefore, Stochastic Return guarantees a positive allocation surplus at return, and a positive minimum revenue for the seller. However, in this section, we show inducing partial learning is always sub-optimal. In other words, Stochastic Return is sub-optimal.

Theorem 1. *If $\mu_0 \in [\underline{\mu}, \overline{\mu}]$, Stochastic return is dominated by either Learning Deterrence or the optimal Free Return. That is,*

$$\Pi^*(\mu_0) = \max\{\Pi^{D}(\mu_0), \Pi^{F}(\mu_0)\}.$$

Sketch of proof. The proof of this theorem proceeds in three steps.

In step 0, we characterize the refund mechanisms implementing the interior learning outcomes (Lemma 2). Specifically, for a fixed s, the return policy $(x_r(\mu, s), t_r(\mu, s))$ as a function of buyer's stopping belief μ induces the buyer to optimally stop learning at belief μ. Then the seller maximizes the expected revenue–a weighted average between the price and the return transfer–over the inducible set of (s, μ). We can formulate the seller's optimization problem for encouraging learning as below.

$$\Pi^{\mathcal{P}}(\mu_0) := \max_{s \in [q^{-1}(\mu_0), Q^{-1}(\mu_0)]} \left\{ \max_{\mu} \; \Pi(\mu, s) = \frac{\mu_0 - \mu}{1 - \mu}(v_h - s) + \frac{1 - \mu_0}{1 - \mu} t_r(\mu, s) \right\} \quad (\mathcal{P})$$
$$\text{s.t.} \quad q(s) \leq \mu \leq Q(s)$$
$$\mu \leq \mu_0$$

In particular, the seller first choose an s that can encourage the type-μ_0 buyer to learn, and then she optimizes the expected revenue over the set of inducible return beliefs. In the next two steps, we show that $\Pi^{\mathcal{P}}(\mu_0) \leq \max\{\Pi^{D}(\mu_0), \Pi^{F}(\mu_0)\}$ with strict inequality for some values of prior belief.

In step 1, we study the inner maximization of $(\mathcal{P})$ to derive the optimal stopping belief for a fixed s (Lemma 3). We characterize the domain of s such that the optimal stopping belief $\mu^*(s)$ is an interior solution, i.e., $\mu^*(s) \in (q(s), Q(s))$. We show that $\mu^*(s)$ is independent of μ_0 and is strictly increasing in s. The solid red curve in Figure 2 depicts $\mu^*(s)$. For simplification, we take $\mu_0 = 0.5$ while illustrating the main idea of this theorem. In this case, the lower boundary point of $\mu^*(s)$ is induced by a Free Return mechanism. While the upper boundary point can be induced by either Stochastic Return or Learning Deterrence, we show Learning Deterrence *strictly* dominates Stochastic Return at this point.

In step 2, we study the outer maximization and show that the seller's revenue is *quasi-convex* along the solid red curve, $\mu^*(s)$. Thus, we establish the sub-optimality of partial learning. The revenue-maximizing mechanism is either Learning Deterrence or Free Return. In the following subsections, we discuss these steps in detail.

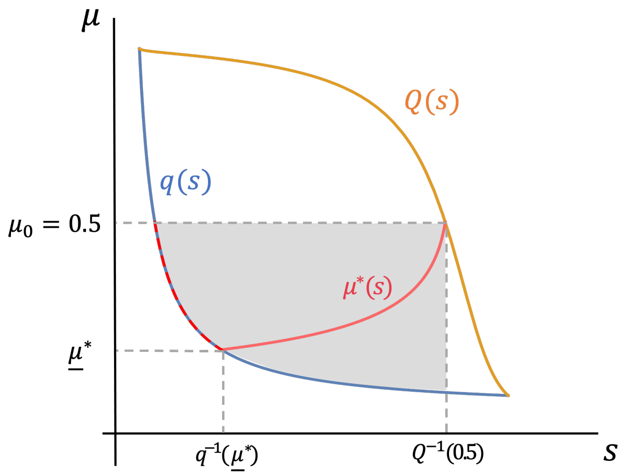


Figure 2: Stochastic Return and interior solutions

5.1 Incentive Compatible Mechanisms

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

$$x_r(\mu, s) = \frac{V_1(\mu, s)}{v_h - v_l}, \tag{4}$$

$$t_r(\mu, s) = \mathbb{E}(v|\mu)\frac{V_1(\mu, s)}{v_h - v_l} - V(\mu, s). \tag{5}$$

Furthermore, the return transfer $t_r(\mu, s)$ increases with both μ and s and with cross derivative equal to 0; and $x_r(\mu, s)$ increases with μ.

Equations (4) and (5) are the familiar smooth pasting and value matching conditions for the buyer to optimally stop at μ given the return policy. In Figure 3, a type-μ buyer obtains expected utility $\mathbb{E}(v|\mu)x_r - t_r$ if he requests a return, whereas he attains 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. One can view the mechanism $\{v_h - s, (x_r(\mu, s), t_r(\mu, s))\}$ as a direct mechanism to screen an interim-type buyer, such that (1) the buyer stops acquiring information if his posterior reaches μ or 1; and (2) the buyer strictly prefers to truthfully report his posterior beliefs.

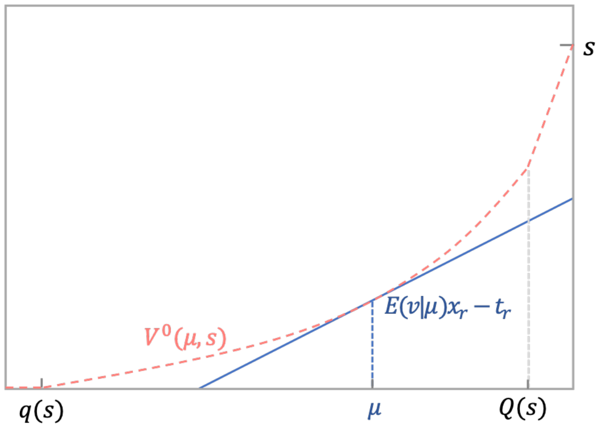


Figure 3: Partial Learning with Stochastic Return

Interestingly, Lemma 2 shows that $t_r(\mu, s)$ increases with μ, meaning that the seller actually obtains a larger return transfer if she enforces earlier stopping, even though she has to compensate the buyer with more opportunity information rent. The main reason is that the seller uses both x_r and t_r towards the buyer's compensation. In particular, when the seller intends to induce a higher stopping belief, she allows the buyer to keep the item with greater probability upon return, thereby inducing the buyer to make a larger return transfer. With similar reasoning, the return transfer $t_r(\mu, s)$ also increases with s. It means that, for a fixed stopping belief, the seller obtains a larger return transfer by charging a smaller selling price.

Figure 4 plots the return transfer for a fixed s. For fixed s, the domain of $t_r(\cdot, s)$ is $[q(s), Q(s)]$. Note that,

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

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

That is, for fixed s, Free Return is the left limit of Stochastic Return. In contrast, the right limit of Stochastic Return is strictly dominated by Learning Deterrence in terms of seller revenue. This follows because the return transfer $t_r(Q(s), s)$ must be smaller than the selling price, $v_h - s$, which implies $t_r(\mu, Q^{-1}(\mu)) < v_h - Q^{-1}(\mu) = t^D(\mu)$. Regarding how much information the buyer acquires, Free Return and Learning Deterrence can be interpreted as opposite limits of Stochastic Return.

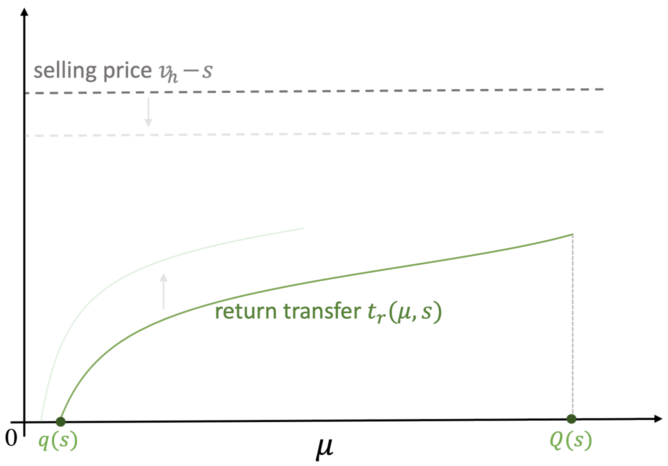


Figure 4: Return transfer

5.2 Inner Maximization

For fixed s, and ignoring the constraint $\mu \leq \mu_0$, consider the internal maximization problem of ($\mathcal{P}$):

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

We readily verify that $\Pi(\cdot, s)$ is quasi-concave on $[q(s), Q(s)]$, implying it attains a maximum either at the boundaries (which reduces to no learning or full learning), or at an interior solution characterized by the first-order condition (which corresponds to partial learning). Denote $\mu^*(s)$ as the maximizer of ($\mathcal{R}$) when it is the solution to the first-order condition. Formally,

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

Recall that $\Pi(\mu, s)$ is a weighted average between the return transfer and the price. Rearranging the first order condition, we obtain[11]

$$\underbrace{\Pr(\text{return}) \frac{\partial t_r(\mu, s)}{\partial \mu}}_{\text{larger return transfer}} = \underbrace{[v_h - s - t_r(\mu, s)] \frac{d\Pr(\text{return})}{d\mu}}_{\text{more frequent return}}. \tag{6}$$

Recall that the seller can gain a larger return transfer $t_r(\mu, s)$ if the buyer stops learning and returns the product at a higher belief μ. However, raising the stopping belief μ increases the probability of receiving a refund request, reducing her revenue from $t_b = v_h - s$ to $t_r(\mu, s)$. We can verify that $\mu^*(s)$ is independent of the prior belief μ_0.

[11] Denote $\Pr(\text{return}) = \frac{1-\mu_0}{1-\mu}$ as the ex-ante probability of return.

Note that $\mu^*(s)$ is strictly increasing in s, shown as the red solid curve in Figure 2. To see this, consider the above trade-off in equation (6) again. Recall that Lemma 2 establishes that the return transfer $t_r(\mu, s)$ increases with s. Therefore, the refund $v_h - s - t_r(\mu, s)$ becomes smaller when s is higher. The seller then cares less about the return rate, and her incentive to gain a larger return transfer is relatively stronger. Thus, she optimally adapts to gain a larger return transfer by inducing a larger stopping belief, meaning that $\mu^*(s)$ increases with s. Furthermore, if s becomes sufficiently high, the refund becomes sufficiently small that gaining a larger return transfer becomes the seller's dominant incentive. She then prefers to induce the maximal stopping belief, rendering the upper boundary $Q(s)$ the optimal return belief. Conversely, if s is sufficiently small, the dominant incentive is to reduce the return rate, and the seller induces the minimal stopping belief, rendering the lower boundary $q(s)$ the optimal stopping belief. Hence, in the last two scenarios, the seller optimally induces no learning and full learning, respectively.

Lemma 3. *Let $\underline{\mu}^*$ be the solution of $\Pi_1(\mu, q^{-1}(\mu)) = 0$. Then $\underline{\mu}^* < 0.5$.*
(1) If $s \leq q^{-1}(\underline{\mu}^*)$, full learning with return belief $q(s)$ is optimal;
(2) If $s \in (q^{-1}(\underline{\mu}^*), Q^{-1}(0.5))$, partial learning with return belief $\mu^*(s)$ is optimal;
(3) If $s \geq Q^{-1}(0.5)$, no learning is optimal.

Lemma 3 summarizes the optimal stopping belief as s varies. The second term of this lemma indicates that partial learning can be seller-optimal if the value of s is intermediate. Given $\mu^*(s)$ being strictly increasing in s, this means that partial learning at stopping belief μ can be induced with optimality only when $\mu \in (\underline{\mu}^*, 0.5)$, shown as in Figure 2. The value 0.5 comes from the observation that the first-order equation, $\Pi_1(\mu, Q^{-1}(\mu)) = 0$, has a unique solution at $\mu = 0.5$.

5.3 Outer Maximization

If the solution to ($\mathcal{P}$) turns out to be interior, it has to be located on the interior path of $\mu^*(s)$. The seller's profit along the path of $\mu^*(s)$ equals:

$$\Pi(\mu, s^*(\mu)) = t_r(\mu, s^*(\mu)) + \frac{\mu_0 - \mu}{1 - \mu}\Big[v_h - s^*(\mu) - t_r(\mu, s^*(\mu))\Big]. \tag{7}$$

where $s^*(\mu)$ represents the inverse of $\mu^*(\cdot)$ for $\mu \in [\underline{\mu}^*, 0.5]$. The ratio $\frac{\mu_0 - \mu}{1 - \mu}$ is the ex-ante probability of a successful sale. Note that the first term is the seller's minimum revenue, and the second term refers to the extra revenue she can obtain if the buyer discovers good news.

We show that seller's profit is quasi-convex along the path of $\mu^*(s)$. Equivalently, $\Pi(\mu, s^*(\mu))$ is quasi-convex on $\mu \in [\underline{\mu}^*, 0.5]$. To see this, suppose that the seller tends to lower the buyer's stopping belief to induce more extended learning, thereby increasing the odds of a successful sale. She can further benefit as the optimal price $v_h - s^*(\mu)$ increases simultaneously. In addition, as the seller adjusts the mechanism to induce a lower stopping belief, the minimum revenue $t_r(\mu, s^*(\mu))$ also decreases,[12] causing the extra revenue gain even more substantial, which reinforces the seller's motive to decrease the buyer's stopping belief. It gives rise to a corner solution at the lower boundary $\underline{\mu}^*$ (see the lower boundary of the solid red curve in Figure 2).

Conversely, suppose the seller tends to induce a higher stopping belief to raise her minimum revenue $t_r(\mu, s^*(\mu))$. She further benefits as the optimal price $v_h - s^*(\mu)$ decreases simultaneously, reinforcing her incentive to increase the stopping belief and raise the minimum revenue. It produces a corner solution at the upper boundary of some feasible regions. When $\mu_0 = 0.5$, the upper boundary is just μ_0 (see the upper boundary of the solid red curve in Figure 2). Thus, the seller's expected revenue $\Pi(\mu, s^*(\mu))$ is quasi-convex in the stopping belief.

As we mentioned previously, inducing full learning corresponds to Free Return, whereas inducing no learning via Stochastic Return is strictly dominated by Learning Deterrence. Therefore, inducing partial learning via Stochastic Return is suboptimal. In other words, if the seller allows a return in the optimal refund mechanism, she requires the buyer to return the product with probability one while issuing a refund. Otherwise, she does not allow a return.

6 Optimal Refund Mechanism

Given Theorem 1, $\max\{\Pi^D(\mu_0), \Pi^F(\mu_0)\}$ determines the value of the optimal mechanism. Let $\gamma = k/\lambda$ be the effective learning cost. Let F be the set of μ_0 such that the seller weakly prefers to choose Free Return. Formally,

$$F = \{\mu_0 \in [\underline{\mu}, \overline{\mu}] : \Pi^F(\mu_0) \geq \Pi^D(\mu_0)\}.$$

[12]Because $t_r(\mu, s)$ increases with both arguments and $s^*(\mu)$ increases with μ.

Theorem 2. *There exists a γ^* such that if $\frac{k}{\lambda} \leq \gamma^*$, then F is a closed interval and $F \subset (v_l/v_h, \bar{\mu})$; if $\frac{k}{\lambda} > \gamma^*$, $F = \varnothing$. The optimal mechanism takes following form:*

1. *No Return (with $t_b = \mathbb{E}(v|\mu_0)$ and $(x_r, t_r) = (1, t_b)$) if $\mu_0 \notin [\underline{\mu}, \overline{\mu}]$;*
2. *Learning Deterrence (with $t_b = t^D(\mu_0)$ and $(x_r, t_r) = (1, t_b)$) if $\mu_0 \in [\underline{\mu}, \overline{\mu}]$ and $\mu_0 \notin F$;*
3. *Free Return (with $t_b = v_h - s^F(\mu_0)$ and $(x_r, t_r) = (0, 0)$) if $\mu_0 \in F$.*

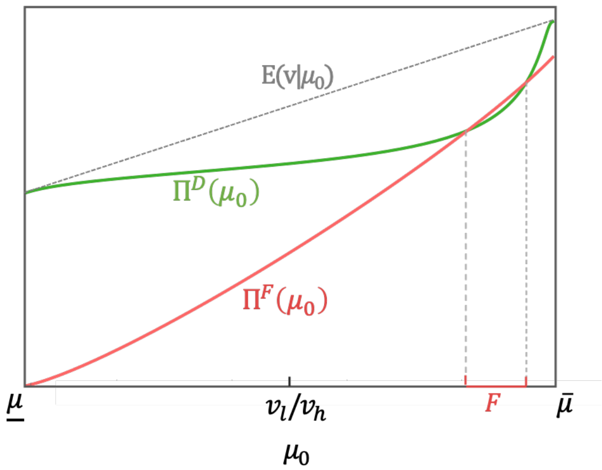


Figure 5: Learning Deterrence and Free Return revenue for small learning cost

Figure 5 depicts the expected revenue of Learning Deterrence (green curve) and the revenue-maximizing Free Return mechanism (red curve) when $\frac{k}{\lambda} < \gamma^*$. These two curves cross twice as shown in the graph. That is, the seller optimally chooses Free Return when the prior belief lies in the red interval F, otherwise, the seller optimally chooses Learning Deterrence.

To interpret this result, intuitively, the buyer's prior belief measures (1) how much the buyer values information ex-ante; and (2) how optimistic the buyer initially is. The gray dotted curve plots the first best allocation surplus $\mathbb{E}(v|\mu_0)$. Recall that if the buyer is very-well informed ex-ante, e.g., $\mu_0 = \underline{\mu}, \bar{\mu}$, the buyer considers learning sub-optimal, therefore the seller can set a non-refundable price equal to the buyer's ex-ante expected valuation to capture the full allocation surplus. In other words, the green curve coincides with the gray dotted curve at the two end points. When the buyer's prior belief becomes less extreme, information becomes valuable. To deter learning, the seller must lower the price to compensate for the buyer's opportunity information rent, which is the difference between the gray dotted line and the green curve. The less extreme the buyer's prior belief, the larger this opportunity information rent. Therefore, as μ_0 moves from either $\underline{\mu}$ or $\overline{\mu}$ toward a more intermediate belief, the seller has to give away a larger amount of allocation surplus to the buyer, rendering Learning Deterrence less profitable. However, if the seller switched to Free Return to encourage learning, she can avoid compensating the buyer's opportunity information rent. That is, instead of significantly decrease the price to deter learning, she can significantly increase the price to encourage learning. Nevertheless, Free Return might induce inefficient trading ex-post, therefore the seller only favors Free Return when the buyer's prior belief is also more optimistic, as it can guarantee a high probability of successful sale. As a result, Free Return is optimal when the buyer's prior belief is less extreme but also more optimistic. The value of v_l/v_h is a measure of optimism as the left boundary of the red interval F can never go below this ratio.

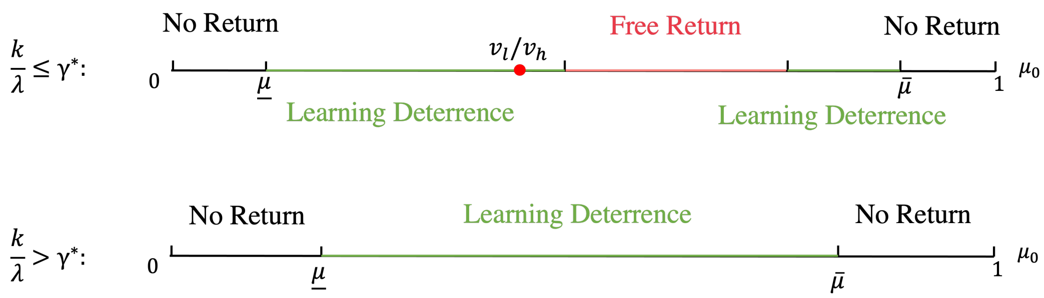
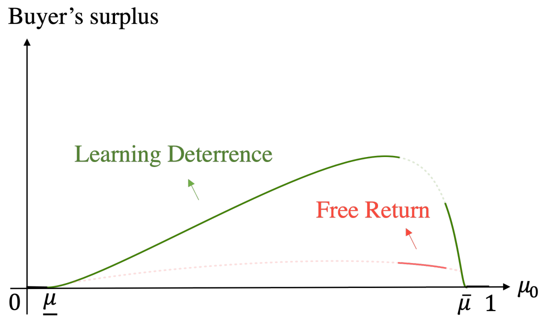


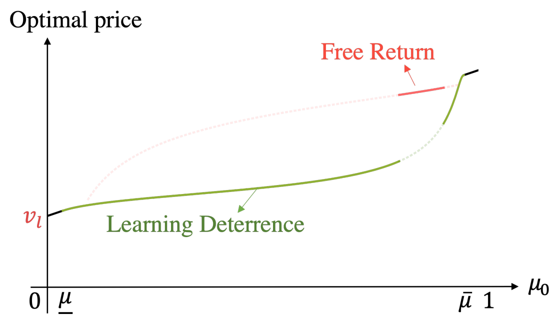
Figure 6: Optimal refund mechanism

When the effective learning cost is large, $\frac{k}{\lambda} > \gamma^*$, the buyer obtains little value from learning. Therefore, the amount of information rent that the seller has to pay the buyer to deter learning is small regardless of his prior belief. Hence, Learning Deterrence becomes more appealing to the seller. Meanwhile, Free Return becomes less profitable as the buyer optimally quits learning earlier, which reduces the ex-ante probability of a successful sale. Therefore, when learning becomes more costly, the set of priors F that supports Free Return as the optimal mechanism shrinks; and eventually becomes an empty set when $\frac{k}{\lambda} > \gamma^*$. See Figure 6.

Though the buyer enjoys a larger information rent if his prior belief is less extreme, he only gains the benefit from it if the seller deters learning. In contrast, if the seller allows Free Return, the buyer then suffers a strict decline in terms of his expected trading surplus (see Figure 7(a)). This is because the seller escalates the price discrimination while encouraging the buyer to learn (see Figure 7(b)). Meanwhile, it is the buyer who takes the cost of learning and inefficient allocation if the seller sets Free Return.



(a) buyer's expected surplus



(b) Optimal selling price

Figure 7: Buyer's surplus and selling price

6.1 Comparative statics

Proposition 4. *The Free Return revenue $\Pi^F(\mu_0)$ is decreasing in the effective learning cost, while the Learning Deterrence revenue $\Pi^D(\mu_0)$ is increasing in the effective learning cost. The set of prior belief supporting Free Return as the optimal mechanism expands if the effective learning cost goes down.*

Figure 8 depicts the seller's revenue, the optimal selling price and the buyer's expected trading surplus while the effective learning cost $\frac{k}{\lambda}$ varies. Interestingly, when learning becomes less costly, the seller optimally allows Free Return more often as deterring learning becomes more expensive for her. This eventually hurts the buyer with relatively optimistic prior.

Proposition 5. (A limit result)

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

As the effective learning cost converges to zero, the buyer can learn almost perfect

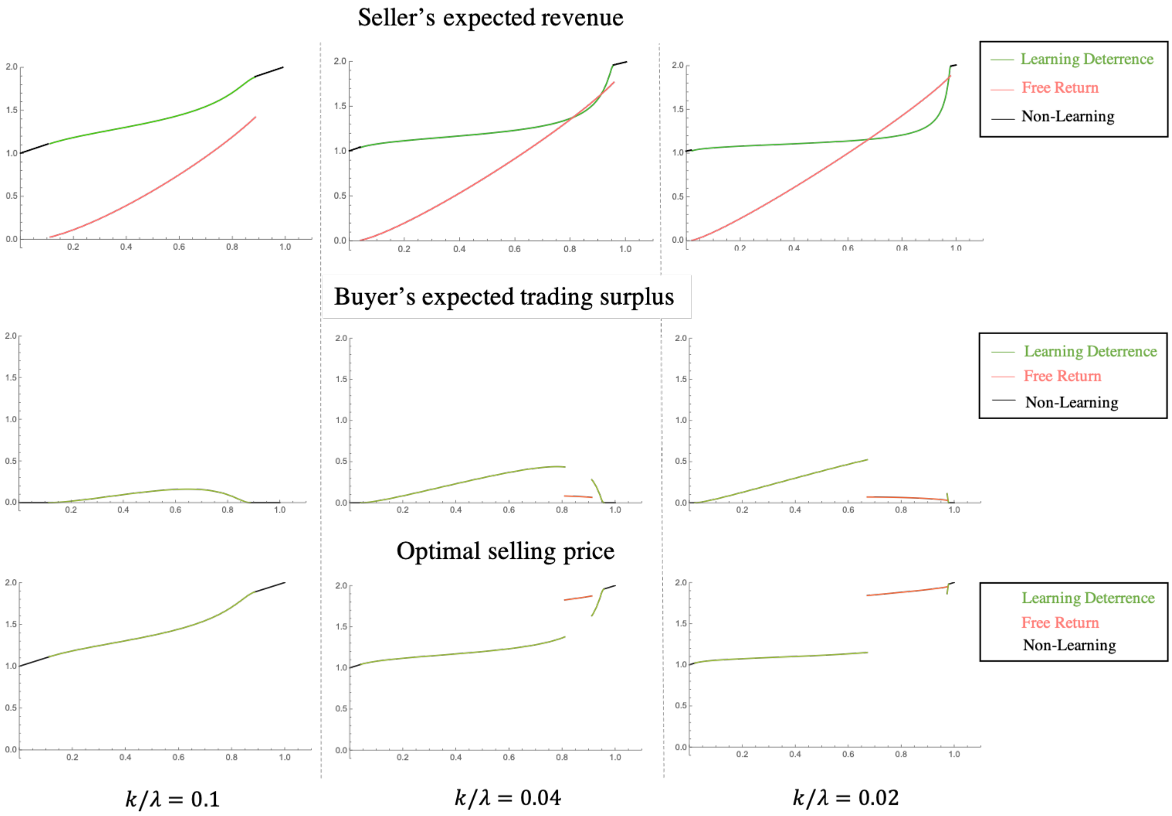


Figure 8: Comparative statics

information. Therefore, with Leaning Deterrence, the seller has to set a price arbitrarily close to v_l, because otherwise, the buyer always has an incentive to learn to avoid consuming the item when his true valuation is low. It relates to the mass market strategy. With Free Return, the seller optimally sets the price arbitrarily close to v_h and lets go of buyers who are almost sure to have a low valuation, which corresponds to the niche market strategy. The ratio $\frac{v_l}{v_h}$ determines the cutoff prior belief at which the seller is indifferent between Free Return and Learning Deterrence. It converges to standard screening result when the buyer privately knows his true valuation.

7 More Efficient Post-purchase Learning

In some scenarios, the learning process is more efficient after a transaction, such as the market of database license, the buyer has access to quite limited information before purchase. Therefore, the transaction itself generates extra information rent.

However, the seller can fully extract this extra information rent by charging a cancellation fee to make the buyer just indifferent between acquiring information before and after purchase. Note that charging a cancellation fee is equivalent to issuing a partial refund, therefore the mechanism space remains the same. Nevertheless, we consider the cancellation fee as a complementary instrument as it is used to extract the additional information rent.

Denote t_u as the cancellation fee. We can then represent the refund mechanism with a cancellation fee as $\{t_b, (x_r, t_r + t_u)\}$. Specifically, if the buyer eventually requests a return, the seller obtains a net return revenue $t_r + t_u$. If the refund mechanism does not allow a return, we let $t_r = t_b$ and $t_u = 0$.

Proposition 6. *If the optimal mechanism allows return, the cancellation fee t_u is the solution to the following equation,*

$$V(\mu_0, s + t_u; \lambda_P) - t_u = V^0(\mu_0, s; \lambda_B). \tag{8}$$

Proposition 6 implies that, for any optimal refund mechanism, the buyer obtains the same ex-ante surplus as if the mechanism prohibited return. If $\lambda_P = \lambda_B$, then $t_u = 0$. Thus, Proposition 6 is a generalization of Lemma 1.

If the optimal mechanism deters buyer learning, then it takes the same form as Learning Deterrence, with price $t_b = t^D(\mu_0; \lambda_B)$ and return policy $(1, t_b)$, regardless of the post-purchase learning rate. If the optimal mechanism encourages learning, then the return policy designed to induce some particular stopping belief μ is obtained in the same way as in Lemma 2. In particular,

$$x_r(\mu, s) = \frac{V_1(\mu, s + t_u; \lambda_P)}{v_h - v_l}, \text{ and } \; t_r(\mu, s) = \mathbb{E}(v|\mu)x_r(\mu, s) - V(\mu, s + t_u; \lambda_P).$$

Thus, to encourage learning, the seller's optimization problem is the following.

$$\max_{s \in [q^{-1}(\mu_0;\lambda_P), Q^{-1}(\mu_0;\lambda_B)]} \left\{ \max_{\mu} \frac{\mu_0 - \mu}{1 - \mu}(v_h - s) + \frac{1 - \mu_0}{1 - \mu}\left(t_r(\mu, s) + t_u\right) \right\} \tag{9}$$

$$\text{s.t.} \quad q(s + t_u; \lambda_P) \le \mu \le Q(s; \lambda_B)$$

$$\mu \le \mu_0$$

If λ_P is close to λ_B,[13] the main result in the benchmark model is robust in the sense that deterministic mechanism is optimal, i.e., $x_r \in \{0, 1\}$, since the cancellation fee

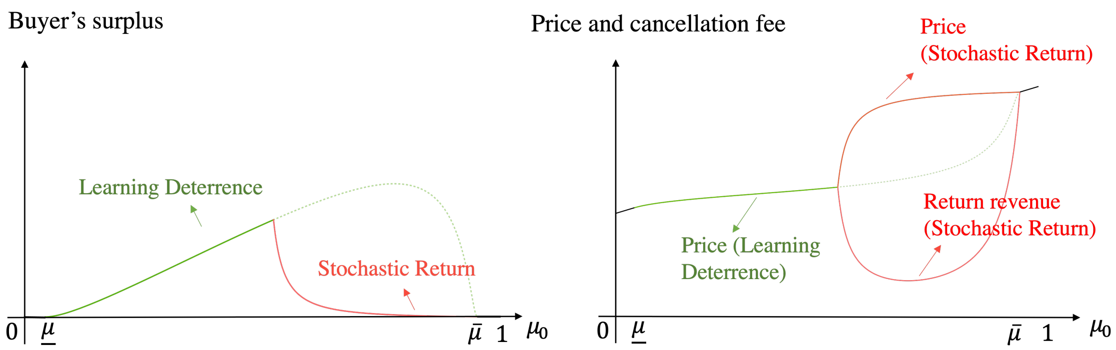
[13] For some prior belief μ_0, the difference between λ_P and λ_B can be very large. For example, let $k = 0.2, \lambda_B = 5, v_h = 2, v_l = 1, \mu_0 = 0.6$. Then inducing partial learning is sub-optimal if $\lambda_P \le 64$.

is not very large. However, $x_r = 0$ implies that there is no trading if the seller matches with a low valuation buyer, which is inefficient as even the low valuation buyer values the product more than the seller. The seller can mitigate this issue if λ_P is sufficiently large.

Proposition 7. *If $\lambda_P \to \infty$ and $\mu_0 \in [\underline{\mu}, \overline{\mu}]$, the optimal refund mechanism takes one of the two forms below:*

1. *Learning Deterrence:*
$$t_b = v_h - Q^{-1}(\mu_0; \lambda_B),\ t_u = 0,\ \text{and } (x_r, t_r) = (1, t_b);$$
2. *Stochastic Return:*
$$t_b = v_h - \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)},\ \text{and } t_u = \frac{k}{\lambda_B(1-\mu_0)}\left(1 + (1-\mu_0)\log\left[\frac{\mu_0 v_h}{\mu_0 v_h - v_l}\right]\right),$$
$$x_r = \frac{k\left(v_h - v_l - (1-\mu_0)\left(v_l + (\mu_0 v_h - v_l)\log\left[\frac{\mu_0 v_h}{\mu_0 v_h - v_l}\right]\right)\right)}{\lambda_B(1-\mu_0)(v_h - v_l)(\mu_0 v_h - v_l)},\ \text{and } t_r = x_r v_l.$$

This proposition discusses the scenario where the buyer can almost learn his true valuation immediately after purchase. Therefore, the buyer consumes the item when his true valuation is high and requests a return if his true valuation is low. In this case, the seller sets a positive allocation rate upon return and sets $t_r = x_r v_l$ to extract the allocation surplus at return. Furthermore, she charges a cancellation fee to extract the extra post-purchase surplus from the buyer.



(a) buyer's expected surplus (b) Price and cancellation fee

Figure 9: Buyer's surplus and optimal selling price if $\lambda_P \to \infty$

In Figure 9, the left panel plots the buyer's ex-ante trading surplus against μ_0 under the optimal refund mechanism for the case $\lambda_P \to \infty$, while the right panel plots the price and the net return revenue $t_r + t_u$ of the optimal refund mechanism. Note that the buyer is still worse off if the seller encourages him to learn due to price discrimination. The difference is that when the seller is indifferent between the two forms of mechanism specified in Proposition 7, the buyer is also indifferent (unlike the discontinuity in Figure 7). This is because the optimal price for the two mechanisms remains the same and the allocation rate x_r of the optimal Stochastic Return is 1.[14] Therefore, there is no efficiency loss, meanwhile, the cost for learning converges to zero, meaning that encouraging the buyer to learn does not impose additional cost to the buyer.

As the seller uses Stochastic Return to mitigate the efficiency loss when a low valuation buyer requests a return, then if $v_l = 0$, such incentive is irrelevant and deterministic mechanism can do equally well.

Corollary 2. *If $\lambda_P \to \infty$ and $v_l = 0$, then the optimal Stochastic Return mechanism in Proposition 7 generates the same expected revenue as this deterministic mechanism—Return with a cancellation fee,*

$$t_b = v_h - \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}, t_u = \frac{k}{\lambda_B(1-\mu_0)}\left(1 + (1-\mu_0)\log\left[\frac{\mu_0 v_h}{\mu_0 v_h - v_l}\right]\right), \text{and } (x_r, t_r) = (0, 0).$$

8 Optimal Mechanism with Bad News

In this section, we consider the opposite learning technology—no news is good news—such that bad news arrives at rate ρ if buyer's true valuation is low (see Keller and Rady (2015)). In this case, the buyer's posterior belief goes up if no news arrives. We call this learning technology *negative learning*. Conversely, we call the good news model as *positive learning*. For simplicity, we assume the learning rate is the same before and after purchase, and let the learning cost k remains the same.

The key difference between positive learning and negative learning is that, under positive learning, buyer returns the product when he becomes sufficiently pessimistic, while under negative learning, the buyer returns the product if he receives bad news which indicates a sure low valuation. Therefore, the seller cannot manipulate the buyer's stopping belief by varying the return policy (x_r, t_r) in bad news model.

[14] x_r is strictly less than one if $\lambda_P < \infty$.

Moreover, if we denote $\eta := v_l x_r - t_r$ as the buyer's surplus while requesting a return upon observing bad news, then under all optimal mechanism,

$$\eta = v_l x_r - t_r = 0.$$

Hence, under negative learning, the seller can only affect the buyer's stopping belief through the selling price t_b, which then determines the buyer's continuation value from learning. For a fixed price, there exists two cutoff beliefs, $g(t_b) \leq G(t_b)$, that determine the buyer's learning behavior. Nevertheless, the lower cutoff $g(t_b)$ becomes the trial belief, which is determined by the indifference between returning the product and continuing to learn,

$$g(t_b) = \{\mu : \mathbb{E}(v|\mu)x_r - t_r = V^N(\mu, t_b)\},$$

where $V^N(\mu, t_b)$ is the buyer's continuation value for learning. The upper stopping belief $G(t_b)$ becomes the consuming belief at which the buyer stops learning and consumes the product. $G(t_b)$ adopts a close form solution in bad news model,

$$G(t_b) = 1 + \frac{k}{\rho(v_l - t_b)}.$$

While varying the selling price, the seller can induce different stopping beliefs. A higher price indicates a higher consuming belief which implies a smaller probability of successful sale.

Denote $G^{-1}(\mu)$ as the inverse function of $G(t_b)$. Thus, the seller can induce a stopping belief of μ if she sets a price equal $G^{-1}(\mu)$. For example, if $t_b = G^{-1}(\mu_0)$, then the seller deters buyer learning. Moreover, let $\bar{G}(\mu_0) := \{\mu : V^N(\mu_0, G^{-1}(\mu)) = 0\}$ be the largest stopping belief that is inducible given the prior belief μ_0. We can then formulate the seller's optimization problem ($\mathcal{N}$) as following.

$$\Pi^{\mathcal{N}}(\mu_0) := \quad \max_{\mu} \quad \Pi^N(\mu) := \frac{\mu_0}{\mu} G^{-1}(\mu) + \frac{\mu - \mu_0}{\mu} t_r \tag{$\mathcal{N}$}$$

$$\text{s.t} \quad \mathbb{E}(v|\mu_0)\frac{t_r}{v_l} - t_r = V^N(\mu_0, G^{-1}(\mu))$$

$$\mu_0 \leq \mu \leq \bar{G}(\mu_0)$$

Similarly, the seller's expected revenue is a weighted average between the selling price and the return transfer. The relative weight depends on both the prior belief and the stopping belief. The seller's revenue equals the deterring learning price at μ_0,[15] i.e., $\Pi^N(\mu_0) = G^{-1}(\mu_0)$, therefore the objective function under negative learning includes the situation of deterring learning, which implies $\Pi^N(\mu_0)$ equals the highest attainable profit. The first constraint comes from $\eta = 0$ and $g(t_b) = \mu_0$, since the lower stopping belief does not enter in the objective function, the seller can always obtain a higher return transfer by increasing the lower stopping belief. The second constraint comes from $V^N(\mu_0, t_b) \geq 0$. Denote $\underline{\mu}^N, \overline{\mu}^N$ as the two beliefs at which the lower and the upper stopping beliefs coincide. Let μ^F be the prior belief at which the seller is indifferent between deterring learning and inducing the longest learning,

$$\mu^F = \min\{\mu_0 \in (\underline{\mu}^N, \overline{\mu}^N] : \Pi^N(\mu_0) = \Pi^N(\bar{G}(\mu_0))\}.$$

Proposition 8. *There exists a γ^{**} such that if $k/\rho < \gamma^{**}$, then $\mu^F < \overline{\mu}^N$ and the optimal mechanism takes following form:*

1. *No Return (with $t_b = \mathbb{E}(v|\mu_0)$ and $(x_r, t_r) = (1, t_b)$) if $\mu_0 \notin [\underline{\mu}^N, \overline{\mu}^N]$;*
2. *Learning Deterrence (with $t_b = G^{-1}(\mu_0)$ and $(x_r, t_r) = (1, t_b)$) if $\mu_0 \in [\underline{\mu}^N, \mu^F]$;*
3. *Free Return (with $t_b = G^{-1}(\bar{G}(\mu_0))$ and $(x_r, t_r) = (0, 0)$) if $\mu_0 \in (\mu^F, \overline{\mu}^N]$.*

Otherwise, if $k/\rho \geq \gamma^{**}$, then $\mu^F = \overline{\mu}^N$ and the optimal mechanism induces no learning for all prior belief and takes the form of No Return and Learning Deterrence.

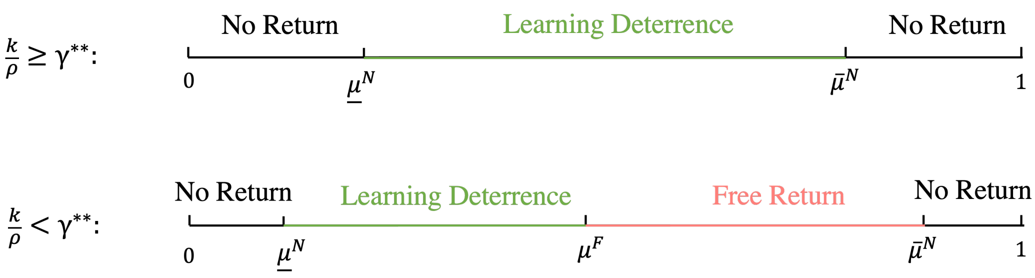


Figure 10: Optimal refund mechanism with bad news

The optimal mechanism under negative learning (described in Proposition 8) takes a similar form as under positive learning (described in Theorem 2). However, the right boundary point of the prior belief that the seller optimally chooses Free Return equals $\overline{\mu}^N$, shown as the second case in Figure 10. Intuitively, if $\mu_0 = \overline{\mu}^N$, the largest

[15] In good news model, the seller's revenue equals the return transfer at μ_0, $\Pi(\mu_0, s) = t_r(\mu_0, s)$, which creates a discontinuity in terms of seller revenue at the prior belief.

inducible stopping belief is just $\bar{G}(\overline{\mu}^N) = \overline{\mu}^N$. Thus the seller's expected revenue from Learning Deterrence is the same as that from Free Return, i.e, $\Pi^N(\mu_0) = \Pi^N(\bar{G}(\mu_0))$ at $\mu_0 = \overline{\mu}^N$.

Note that the optimal Free Return mechanism induces the longest stopping belief $\bar{G}(\mu_0)$ so that the type-μ_0 buyer obtains zero participation value given the definition of $\bar{G}(\mu_0)$. In other words, Free Return further hurts the buyer under negative learning.[16] This is driven by the nature of the learning technology. Specifically, under negative learning, the buyer becomes more optimistic if no news arrives and his continuation value eventually goes up, therefore the seller can keep raising the price until fully capturing the buyer's ex-ante trading surplus. However, under positive learning, the buyer becomes more pessimistic if no news arrives and his continuation value eventually decreases to zero so that he requests a return. Therefore, the seller has to provide the buyer with positive ex-ante surplus to fulfill his participation.

Figure 11 depicts the buyer's participation value against his prior belief under both negative learning and positive learning. Under negative learning, the buyer can only obtain positive participation value if the mechanism deters learning.

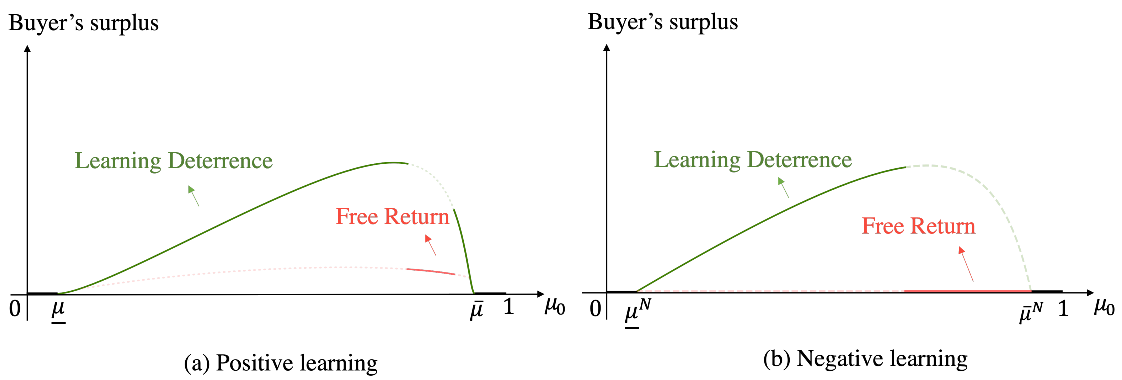


Figure 11: The buyer's ex-ante surplus

9 Discussion

In this section, we discuss several extensions of the baseline model. First, we consider scenarios where the seller cannot freely adjust the selling price. In particular, we

[16]Under positive learning, the buyer can still obtains a positive surplus with Free Return.

focus on the case of an exogenous price or a regulated price cap. We show that Stochastic Return, which induces partial learning, can be optimal in such scenarios. Next, we consider cases in which the seller values the product higher than a low-valuation buyer does. Consequently, the seller intrinsically prefers to encourage learning as return creates efficiency. This shrinks the set of priors that supports Learning Deterrence. Lastly, we argue that the return mechanism we discussed in the baseline model is without loss of generality under a more general framework.

9.1 Optimality of Stochastic Return

Given Lemma 3, Stochastic Return might turn out to be an optimal mechanism in situations in which the seller cannot freely adjust the price, for example, if the price is exogenously determined by other parties. This is the case for some online retail platforms, which can control their return policies but must set prices determined by their suppliers. Other scenarios in which Stochastic Return might be optimal can arise if the prices are driven down by price competition when similar products are sold by multiple sellers, or if a price cap is imposed by the regulator.

Corollary 3. *Suppose that the price is constrained by a price cap* $t^c = v_h - s^c$ *with* $s^c \in (q^{-1}(\underline{\mu}^*), Q^{-1}(0.5))$*. If* $\mu_0 \in (\mu^*(s^c), Q(s^c))$*, the optimal refund mechanism takes one of the two forms below:*

1. *Learning Deterrence with* $t_b = t^D(\mu_0)$ *and* $(x_r, t_r) = (1, t_b)$*;*
2. *Stochastic Return with* $t_b = t^c$ *and* $(x_r, t_r) = (x_r(\mu^*(s^c), s^c), t_r(\mu^*(s^c), s^c))$*.*

Suppose that the price $t_b = v_h - s$ is exogenous with $s \in (q^{-1}(\underline{\mu}^*), Q^{-1}(0.5))$. Lemma 3 implies that for prior belief $\mu_0 \in (\mu^*(s), Q(s))$, the optimal mechanism induces a stopping belief $\mu^*(s)$, and takes the form of Stochastic Return with a price t_b and a return policy $(x_r(\mu^*(s), s), t_r(\mu^*(s), s))$. With a price cap t^c, the seller can adjust the price within the range of values smaller than t^c. As the seller's revenue is quasi-convex in the price when taking into account changes in the optimal stopping belief as the price varies, the optimal mechanism either reduces the price to $t^D(\mu_0)$ to prevent the buyer from private learning, or raises the price to t^c and induces partial learning with the stopping belief $\mu^*(s^c)$.

9.2 Positive seller valuations

In this section, we discuss the situations in which the seller has a positive valuation of the product. This could be, for example, because the seller stands to gain positive revenue from a resale. We use u to denote the seller's product valuation or reservation value. Figure 12 depicts her revenue from Learning Deterrence (green curves) and the optimal Free Return (red curves) as u varies from 0 to v_h. The blue dashed lines represent the values of u. Note that the seller's revenue from Learning Deterrence remains constant when u varies, as it equals the full allocation surplus minus the information rent with which she compensates the buyer to deter learning. Both parts solely depend on the buyer's valuation. However, the seller's revenue from Free Return increases with u as she can collect her reservation value if the buyer returns the product.[17] Note that Free Return and Learning Deterrence are only relevant when the buyer is less well-informed ex ante. In other words, if the prior belief $\mu_0 \notin [\underline{\mu}, \overline{\mu}]$, the buyer deems learning sub-optimal and the seller can set a price equal to the expected buyer valuation and capture the full trading surplus (the black curves).

First, consider the cases where $u \leq v_l$. In this case, efficiency requires trading with probability one, and the seller can always set a price $t_b = v_l$ to prevent the buyer from learning and obtain a revenue higher than her reservation value. That is, the blue line is certainly to lie below the green curve and the black curves. Therefore, if $\mu_0 \in [\underline{\mu}, \bar{\mu}]$, the revenue from the optimal mechanism is determined by the maximum revenue that lies between Learning Deterrence and Free Return. Meanwhile, as larger u induces a greater Free Return revenue, the set of prior beliefs at which the seller optimally chooses Free Return expands. Otherwise, for extreme prior beliefs, the revenue is depicted by the black curves.[18]

Next, we discuss the cases where $u > v_l$. If $\mu_0 \in [\underline{\mu}, \bar{\mu}]$ and the prior belief is low (e.g., $\mathbb{E}(v|\mu_0) < u$), then deterring learning to induce immediate trading becomes a dominated strategy, as keeping the item provides the seller with a higher payoff (the blue line is higher than the green curve in some regions of each plot). Thus, the

[17] With a positive u, the objective function for Free Return becomes to: $\max_s \frac{\mu_0 - q(s)}{1-q(s)}(v_h - u - s) + u$.

[18] Alternatively, we can normalize the buyer's valuation by subtracting the seller's valuation. That is, if $u \leq v_l$, the optimal mechanism is characterized by Theorem 2 after replacing the buyer valuation with $v_h - u > v_l - u \geq 0$ and setting the seller's valuation to 0.

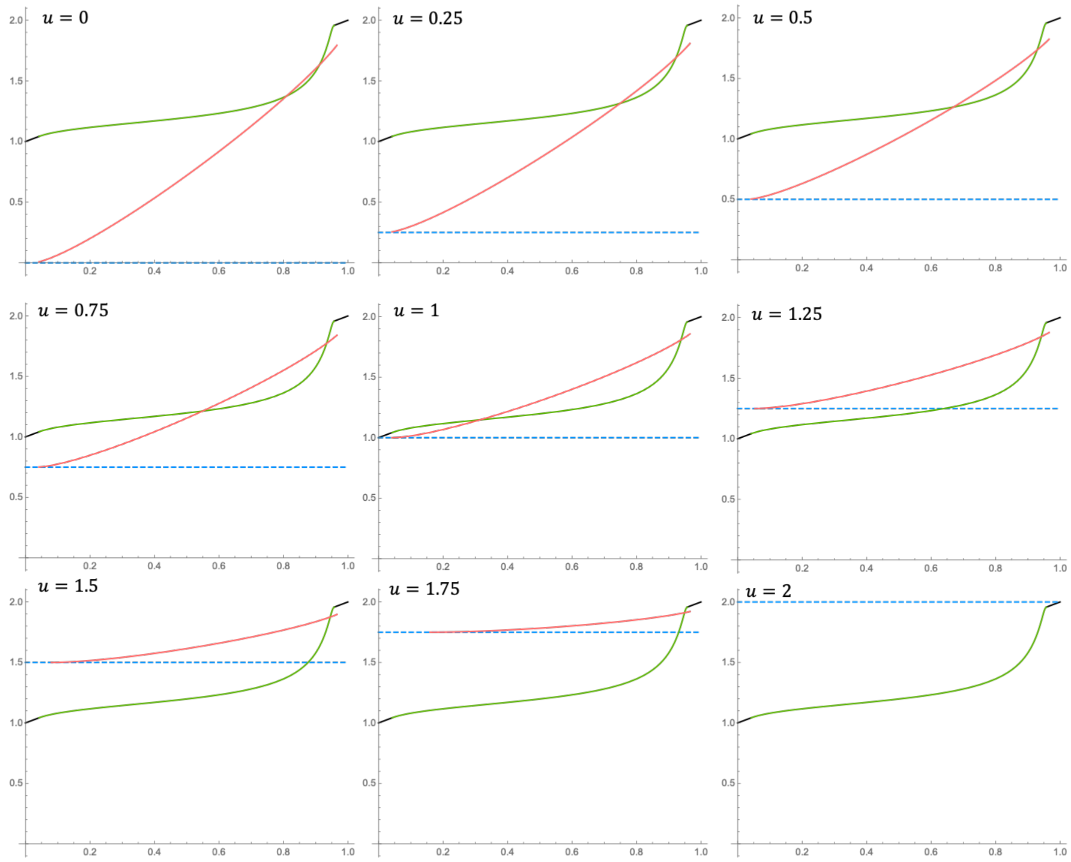


Figure 12: Simulation for different seller valuations ($v_l = 1, v_h = 2, \gamma = 0.04$)

seller has an even stronger incentive to encourage buyer learning, as a return would create efficiency. As a consequence, the set of prior beliefs at which Free Return dominates Learning Deterrence expands as u increases. Meanwhile, with a larger u, the cutoff belief $q(v_h - u)$ above which the seller strictly prefers Free Return rather than keeping the item also increases.[19] Above all, the set of prior beliefs at which the seller optimally chooses Learning Deterrence shrinks. However, the set of prior beliefs that supports Free Return shifts to the right and may eventually vanish as keeping the item becomes more attractive. For example, when $u = v_h$, the seller never sells.

9.3 A more general framework

Notice that a refund mechanism $\{t_b, (x_r, t_r)\}$ can actually be interpreted as a binary menu $\{(x_b = 1, t_b), (x_r, t_r)\}$. Essentially, the seller commits to such a binary

[19] If $\mu_0 = q(v_h - u)$, the optimal pricing for free return equals u.

menu to screen the buyer who is initially uninformed but observes different information outcomes after private learning. In principle, the seller can design any arbitrary menu containing arbitrary numbers of allocation-transfer pairs. With our information process, it is without loss of generality to assume such a binary menu $\{(1, t_b), (x_r, t_r)\}$.[20] More concretely, the buyer observes the binary menu and decides how much information to acquire. Once he finds that the payoff from accepting either option within the menu weakly dominates the continuation value from further learning, he stops and chooses optimally between the two options. Anticipating this, the seller chooses among different binary menus to maximize her expected revenue. Within this framework, our results still hold.

To further elaborate, consider a standard adverse selection model where the agent's private type (valuation) is supported on a closed interval. The revelation principle implies that the principal designs a menu that incentivizes the agent to truthfully reveal his type. In other words, there is a one-to-one mapping from the agent types to the menu options. Here, designing a menu to screen buyers with different posterior beliefs requires more constraints. First, the interim incentive constraints require buyers with different posterior beliefs to be willing to truthfully reveal these beliefs. Second, the buyer must be willing to stop learning when his posterior belief reaches either of the two posteriors for which the binary menu is designed (this is referred to as "implementability" in Mensch (2020)). In our paper, by assuming exponential experimentation, these two sets of constraints are directly implied by the buyer's optimality to stop learning at some particular posterior belief. That is, we only require (x_r, t_r) to satisfy the well-known smooth pasting and value matching conditions at that particular posterior belief (see Lemma 2). Furthermore, given the buyer's optimality of stopping, the first sets of constraints (interim incentive constraints) are always slack; otherwise, learning would not be necessary.

Our reason for distinguishing between $(1, t_b)$ and (x_r, t_r) is that they play very different roles in shaping the buyer's learning behavior. Specifically, the selling price,

[20]The optimality of such a binary menu is implied by Lemmas 1 and 2. For fixed t_b, the seller is maximizing her expected revenue over the distribution of the buyer's posterior beliefs. Given a binary state space, standard concavification implies the optimality of inducing binary posterior beliefs. Mensch (2020) also claims the optimality of binary menus under binary state space, but without requiring $x_b = 1$, as he allows the buyer to acquire flexible information with a posterior-separable information cost.

$(1, t_b)$, determines the net surplus that the buyer obtains if a Poisson jump occurs, and thereby determines the continuation value he can attain through learning. Meanwhile, the return policy, (x_r, t_r), is designed to truncate the buyer's sequential learning, which allows the seller to induce more flexible buyer learning while keeping the price constant.

10 Conclusion

This paper discusses the seller's optimal refund mechanism when interacting with a buyer who can privately acquire information before and after purchase. A refund mechanism is essentially an option contract creating the buyer an option value, which further affects the buyer's learning outcomes. When the difference between before-purchase learning and post-purchase learning is not very large, the optimal refund mechanism is a deterministic mechanism and it either induces full learning or deters the buyer from private learning. However, if the information attainable before purchase is limited and the post-purchase learning is extremely informative, the seller uses stochastic mechanism to increase trading efficiency and charges a cancellation fee to extract the surplus from the buyer.

Appendix

Proposition 1

Proof. We prove this proposition by verifying $Q(s) \geq q(s)$ if $s \in [\underline{s}, \bar{s}]$, while the equality holds at $\underline{s}$ and $\bar{s}$. Recall that $Q(s) = \{\mu : V(\mu, s) = \mathbb{E}(v|\mu) - (v_h - s)\}$. By setting $s = D(\mu) := Q^{-1}(\mu)$, the type-$\mu$ buyer is indifferent between accepting the price and exerting learning. Let $\tilde{\mu}(\mu) := q(D(\mu))$ be the quitting belief if $s = D(\mu)$.

Claim 1. *The domain of $\tilde{\mu}(\mu)$ is $[\underline{\mu}, \overline{\mu}]$. $\tilde{\mu}(\mu) \leq \mu$ and the equality holds only at the two end points. $\tilde{\mu}(\mu)$ is increasing and symmetric about the line 1-μ. $\mu - \tilde{\mu}(\mu)$ increases first and then decreases in μ.*

Proof. Recall the definition of $D(\mu)$,

$$V(\mu, D(\mu)) = \mathbb{E}(v|\mu) - (v_h - D(\mu)). \tag{10}$$

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

$$\frac{dD(\mu)}{d\mu} = \frac{k(k - \lambda D(\mu))}{\lambda^2(1-\mu)^2\mu D(\mu)} = \frac{k[\tilde{\mu} - 1]}{\lambda(1-\mu)^2\mu} < 0. \tag{11}$$

Besides,

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

Thus, $\tilde{\mu}(\mu)$ is a differential equation with initial point $(\underline{\mu}, \underline{\mu})$,[21] and its solution is,[22]

$$-\frac{1}{\tilde{\mu}} - \log[1-\tilde{\mu}] + \log[\tilde{\mu}] = \frac{1}{1-\mu} - \log[1-\mu] + \log[\mu] - \frac{\lambda(v_h - v_l)}{k}. \tag{12}$$

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 are smaller then 0.5 and $f'(\cdot) < g'(\cdot)$ when both arguments are larger then 0.5.[23] Therefore $f(\cdot)$ and $g(\cdot)$ cross only at the two boundary points and therefore

[21] To verify $(\underline{\mu}, \underline{\mu})$ is an initial point. Recall $\underline{\mu} = q(\bar{s})$ and the binding Learning-Feasibility constraint implying $\mathbb{E}(v|\underline{\mu}) - (v_h - \bar{s}) = 0$. Meanwhile $V(q(\bar{s}), \bar{s}) = V(\underline{\mu}, \bar{s}) = 0$. Given equation (10), $D(\underline{\mu}) = \bar{s}$. Thus, $\tilde{\mu}(\underline{\mu}) = q(D(\underline{\mu})) = \underline{\mu}$.

[22] 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.

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

$\tilde{\mu}(\mu) < \mu$ 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 $(1-\tilde{\mu}, 1-\mu)$. It is easy to verify that, if equation (12) holds at a point $(\mu, \tilde{\mu})$, then equation (12) still holds at the reflection point $(1-\tilde{\mu}, 1-\mu)$. Now, we want to show that $\mu - \tilde{\mu}(\mu)$ is single-peaked, increasing first and then decreasing in μ. 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 $\tilde{\mu}(\mu) = 1-\mu$.[24] As $\tilde{\mu}(\mu)$ is increasing in μ and symmetric about $1-\mu$, it follows that $\tilde{\mu}'(\mu) = 1$ has a unique interior solution. $\square$

When $s \notin [\underline{s}, \bar{s}]$, as the Learning-Feasibility constraint fails, no learning is optimal. If $s \in [\underline{s}, \bar{s}]$, note that $Q(D(\mu)) - q(D(\mu)) = \mu - \tilde{\mu}(\mu)$. Taking the derivative with respect to μ yields $(Q' - q')D' = 1 - \tilde{\mu}'$. Because $D'(\mu) < 0$, $Q'(s) - q'(s)$ is positive for small s and then negative for large s, and $Q(s) = q(s)$ at $\underline{s}$ and $\bar{s}$.[25] The difference, $Q(s) - q(s)$, is single-peaked in s. That is, for all $s \in [\underline{s}, \bar{s}]$, $Q(s) \geq q(s)$ with equality holding at the two end points. Then it is easy to verify $V(\mu, s) \geq \max\{0, \mathbb{E}(v|\mu) - (v_h - s)\}$ if $\mu \in [q(s), Q(s)]$. Then the construction of Proposition 1 is optimal based on the standard arguments in the exponential experimentation. $\square$

Lemma 1

Proof. To simplify the exposition, we omit the notion of s in the buyer's value function, as the lemma is true for any fixed s. Let $V_B(\mu(\tau))$ be the buyer's value function for pre-purchase learning (enter-the-market value). It is characterized by the Bellman equation below:

$$V_B(\mu(\tau)) = \max\{0, V_P(\mu(\tau)), -kd\tau + \mu(\tau)\lambda d\tau s + (1-\mu(\tau)\lambda d\tau)V_B(\mu(\tau+d\tau))\}. \quad (13)$$

Different from the Bellman equation for No return, if the buyer stops learning by purchasing the item, he obtains the purchase value $V_P(\mu(\tau))$ instead of the consumption value $\mathbb{E}(v|\mu) - (v_h - s)$, as he might also learn after purchase if a return

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

[25] Recall that $D(\underline{\mu}) = \bar{s}$ and $D(\bar{\mu}) = \underline{s}$ by Learning-Feasibility.

is allowed. The purchase value $V_P(\mu(\tau))$ is characterized as below:

$$
\begin{aligned}
V_P(\mu(\tau)) = \max\{&\mathbb{E}(v|\mu(\tau)) - (v_h - s), \mathbb{E}(v|\mu(\tau))x_r - t_r, \\
&- kd\tau + \mu(\tau)\lambda d\tau s + (1 - \mu(\tau)\lambda d\tau)V_P(\mu(\tau + d\tau))\}.
\end{aligned} \tag{14}
$$

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 (13), (14), and (2) lead to the same differential equation (ODE). Obviously, $V_B(\mu) \geq V_P(\mu)$ and $V_B(\mu) \geq V^0(\mu)$.

To show this lemma, we prove the following equality:

$$
\max\{V_B(\mu), V_P(\mu)\} = V^0(\mu), \ \forall \mu. \tag{15}
$$

Suppose the seller intends to set a harsh return policy with which, if the buyer purchases the item and stops post-purchase learning at belief μ,[26] he obtains payoff $V_P(\mu) < V_B(\mu)$. Then, a rational buyer could simply not purchase the item and perform pre-purchase learning, which implies $V_B(\mu) = V^0(\mu)$ and equality (15) holds .

Moreover, suppose the seller instead offers a benevolent return policy intending to reward the buyer for purchasing the item early on. Under this policy, the buyer purchases the item at some point and, while he stops post-purchase 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 allocation surplus minus the buyer's payoff:[27]

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

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

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

Slope $V_P'(\mu)$ and 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

[26] Since the buyer always stops learning if his belief jumps to 1, our use of the term stopping belief refers to the non-degenerate stopping beliefs.

[27] This is implied by the optimality (known by smooth-pasting) to stop learning at μ. Formally, see Lemma 2.

the above differential equation) to increase the return transfer and in the meantime preserve the same buyer's optimal stopping rule,[28] which is a contradiction of optimality, and we obtain condition (15). □

Corollary 1

Proof. Given Lemma 1, for all optimal refund mechanisms, $V_B(\mu_0, s) = V^0(\mu_0, s)$. Given Proposition 1, if $\mu \notin [\underline{\mu}, \bar{\mu}]$, then $V_B(\mu_0, s) = \max\{0, \mathbb{E}(v|\mu_0) - (v_h - s)\}$ and the buyer does not perform learning regardless of s. Suppose the seller offers a "No Return" mechanism with price $t_b = \mathbb{E}(v|\mu_0)$. Then, the buyer is indifferent between purchasing and quitting, with the seller-preferred tie breaking rule, the seller obtains a revenue equal to $\mathbb{E}(v|\mu_0)$. Note that the seller's profit equals the joint surplus minus the surplus that buyer obtains from the trade. In this case, the joint surplus attains the full allocation surplus and the buyer gets zero trading surplus. That is to say, the mechanism $\{\mathbb{E}(v|\mu_0), (1, \mathbb{E}(v|\mu_0))\}$ is optimal. □

Proposition 2

Proof. With Learning Deterrence, $s = D(\mu_0) := Q^{-1}(\mu_0)$, trading happens with probability one and therefore the joint surplus attains the full allocation surplus $\mathbb{E}(v|\mu_0)$.

First, we prove the first term. Proposition 1 and the Learning-Feasibility constraints imply $V(\mu, D(\mu)) = 0$ at $\underline{\mu}$ and $\bar{\mu}$. Rearranging equation (10) gives:

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

Taking derivative w.r.t μ and plugging in equation (11) gives:

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

where $A = \frac{k}{\lambda(v_h - v_l)} = (1 - \underline{\mu})\underline{\mu} \in (0, \frac{1}{4})$.[29] It is easy to verify $\frac{dV(\mu,D(\mu))}{d\mu} = 0$ at $\underline{\mu}$ or $\overline{\mu}$. To prove that $V(\mu, D(\mu))$ is single-peaked in μ, we only need to show

[28] By inducing the same stopping beliefs, the ex-ante probabilities of return and successful sale are the same regardless of when the buyer purchases the item, i.e., switches to post-purchase learning.

[29] 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, $A \in (0, \frac{1}{4})$.

that $\frac{dV(\mu,D(\mu))}{d\mu}=0$ has a unique solution when $\mu\in(\underline{\mu},\overline{\mu})$, as $V(\mu,D(\mu))>0$ when $\mu\in(\underline{\mu},\overline{\mu})$. That is, the two equations below have a unique solution when $\mu\in(\underline{\mu},\overline{\mu})$, as $\tilde{\mu}$ is the implicit solution of (12).

$$-A\frac{(1-\tilde{\mu})}{(1-\mu)^2\mu}+1=0 \tag{16}$$

$$-\frac{1}{\tilde{\mu}}+\log\left[\frac{\tilde{\mu}}{1-\tilde{\mu}}\right]=\frac{1}{1-\mu}+\log\left[\frac{\mu}{1-\mu}\right]-\frac{1}{A} \tag{17}$$

Substituting equation (16) into (17), we have,

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

Denote the LHS as $h(\mu)$. 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 that $h(\mu)$ first decreases and then increases and then decreases again on $[\underline{\mu},\overline{\mu}]$. Taking the derivative of $h(\mu)$ w.r.t μ gives:

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

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 $\mu<1/3$, 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 at most 4 roots: $1/3, 1$, and at most two roots from $(1-\mu)^2\mu-A=0$.[30] We can show that $z(\mu)$ crosses $y(\mu)$ twice in the support $[\underline{\mu},\overline{\mu}]$, first from above and then from below.[31]

Next, the monotonicity of $t^D(\mu_0)=v_h-D(\mu_0)$ can be directly obtained from (11). 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 $t^D(\overline{\mu})=\mathbb{E}(v|\overline{\mu})$. □

[30] $z'(\mu)=2[(1-\mu)^2\mu-A](3\mu-1)(\mu-1)$. 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 $\mu=\underline{\mu},\overline{\mu}$.

[31] (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 $A\in(4/27,1/4)$, and hence we have the same double crossing given $y(\underline{\mu})<z(\underline{\mu})$ and $y(\overline{\mu})<z(\overline{\mu})$.

Propositions 3

Proof. We prove these two propositions together. First, we solve the explicit solution for $s^F(\mu_0)$ and $\Pi^F(\mu_0)$. Denote $\Pi(q(s), s) = \frac{\mu_0 - q(s)}{1-q(s)}(v_h - s)$ as the objective function of ($\mathcal{F}$), and it is concave given that the second order total derivative w.r.t s is negative.[32] Thus, the maximizer is pinned down by the first order condition, which leads to be solution below:

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

and

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

Taking the derivative of $s^F(\mu_0)$ w.r.t μ_0 gives:

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

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

Next, we need to verify that, if $\Pi^{\mathcal{F}}(\mu_0) \geq \Pi^D(\mu_0) = t^D(\mu_0)$, then $q^{-1}(\mu_0) \leq s^F(\mu_0) \leq Q^{-1}(\mu_0)$. It is obvious that, if $\Pi^{\mathcal{F}}(\mu_0) \geq t^D(\mu_0)$, then $v_h - s^F(\mu_0) > t^D(\mu_0) = v_h - Q^{-1}(\mu_0)$, as the expected probability of a successful sale is less than one with Free Return. Hence, $s^F(\mu_0) \leq Q^{-1}(\mu_0)$ holds trivially. To show $q(s^F(\mu_0)) < \mu_0$, we plug in the explicit expression of $s^F(\mu_0)$ and obtain, $\sqrt{\frac{\mu_0}{1-\mu_0}} > \sqrt{\frac{k/(\lambda v_h)}{1-k/(\lambda v_h)}}$. This inequality is true because $\frac{k}{\lambda v_h} < \underline{\mu} < \mu_0$. □

Lemma 2

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 learning at a belief μ different from $q(s)$, $V_P(\mu, s)$ must be equal to $V^0(\mu, s)$. Otherwise, the buyer strictly prefers to continue his pre-purchase learning and does not stop. Furthermore, to ensure that it is a best response for the buyer to stop at belief μ given the return policy (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

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

induced stopping belief μ must belong to the set $[q(s), Q(s)]$, in which $V^0(\mu, s) = V(\mu, s)$. That is,

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

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

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

$$t_r(\mu, s) = -\frac{kv_l - \lambda\mu v_l s - k\mu v_h \left[\log\left(\frac{\mu}{1-\mu}\right) - \log\left(\frac{k}{\lambda s - k}\right)\right]}{\lambda\mu(v_h - v_l)}. \tag{18}$$

Taking partial derivative w.r.t μ and s separately gives:

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

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

and the cross derivative is 0. Moreover, as $V(\cdot, s)$ is convex in μ, $x_r(\cdot, s)$—proportional to $V_1(\cdot, s)$—is therefore increasing in μ. □

Lemma 3

Proof. First, we discuss the first-order condition. Explicitly,

$$\Pi_1(\mu, s) = \frac{(1-\mu_0)}{(1-\mu)^2(v_h - v_l)} \underbrace{\left[v_h(-v_h + s + v_l) + \frac{k(\mu(v_h - 2v_l) + v_l)}{\lambda\mu^2} + \frac{kv_h(\log[\frac{\mu}{1-\mu}] - \log[\frac{k}{\lambda s - k}])}{\lambda}\right]}_{\equiv\Upsilon(\mu)}.$$

Since $\mu \in [\underline{\mu}, \bar{\mu}]$, $\Pi_1(\mu, s) = 0$ has the same solution with $\Upsilon(\mu) = 0$.

$$\Upsilon'(\mu) = \frac{k(1-2\mu)\mu v_h + 2k(1-\mu)^2 v_l}{\lambda(\mu - 1)\mu^3}.$$

The numerator of $\Upsilon'(\mu)$ is a well-behaved second-order polynomial, which is verified to have a unique root between 0 and 1, and is larger than 0 at $\mu = 0$, and smaller than 0 at $\mu = 1$. Thus, $\Upsilon'(\mu)$ crosses 0 only once and from below, which implies $\Upsilon(\mu)$ is initially decreasing and then increasing. Therefore, $\Upsilon(\mu)$ has at most two roots in $[0, 1]$, denoted as $\mu^*_-(s) \leq \mu^*_+(s)$. Furthermore, $\Upsilon(\mu)$ is increasing in s. Therefore, the smaller root is the local maximizer of $\Pi(\mu, s)$ which is increasing in s, while the larger root is the local minimizer of $\Pi(\mu, s)$ which is decreasing in s, and if the two roots coincide, $\mu^*_-(s) = \mu^*_+(s) > 0.5$.[33] Thus, if there exists a $\mu^*_+(s)$, it is larger than 0.5.

Let $s^*(\mu) = \{s : \Pi_1(\mu, s) = 0\}$.[34] Given the above argument, it is a single-valued continuous function, which is initially increasing and then decreasing in μ. Furthermore, it is clear that when $\mu \leq 0.5$, $s^*(\mu)$ is increasing. To introduce one more notation, let $\bar{t}_r(\mu) := t_r(\mu, Q^{-1}(\mu))$. It is the envelope of all inducible return transfers. Formally,

$$t_r(\mu, s) \in [0, \bar{t}_r(\mu)] \Longleftrightarrow \mu \in [q(s), Q(s)].$$

To see this, consider the direction from the right to the left first. Recall that $t_r(\mu, s)$ is increasing in both arguments. If $\mu \geq q(s)$, then $t_r(\mu, s) \geq t_r(q(s), s) = 0$; and if $\mu \leq Q(s)$, then $s \leq Q^{-1}(\mu)$ as $Q(s)$ decreases in s, which then implies $t_r(\mu, s) \leq t_r(\mu, Q^{-1}(\mu))$. The opposite direction is trivial.[35]

To prove Lemma 3, we want to show that $\Pi(\mu, s)$ is quasi-concave on $\mu \in [q(s), Q(s)]$. Specifically, we show $t_r(\mu^*_+(s), s) > \bar{t}_r(\mu^*_+(s))$, which then implies $\mu^*_+(s) > Q(s)$. The following claim pins down the set of μ such that $t_r(\mu, s^*(\mu)) \in [0, \bar{t}_r(\mu)]$.

Claim 2. *$\bar{t}_r(\mu)$ with domain $[\underline{\mu}, \overline{\mu}]$ first increases and then decreases in μ. $t_r(\mu, s^*(\mu))$ single crosses $\bar{t}_r(\mu)$ at 0.5 from below, and $\{\mu : t_r(\mu, s^*(\mu)) \in [0, \bar{t}_r(\mu)]\} = [\underline{\mu}^*, 0.5]$.*

Proof. It is obvious that $\bar{t}_r(\mu) \geq 0$ when $\mu \in [\underline{\mu}, \overline{\mu}]$, with equality hold at the two end points. Recall that $D(\mu) := Q^{-1}(\mu)$. Taking derivative of $\bar{t}_r(\mu)$ w.r.t μ gives

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

The term in square brackets is decreasing. It's positive when $\mu = \tilde{\mu}(\mu) = \underline{\mu}$, and negative when $\mu = \tilde{\mu}(\mu) = \overline{\mu}$. Hence, $\bar{t}_r(\mu)$ is increasing first and then decreasing. Next, we show that $t_r(\mu, s^*(\mu)) = \bar{t}_r(\mu)$ has a unique solution of 0.5. Since $t_r(\mu, s)$ is increasing in s, to find the solution of $t_r(\mu, s^*(\mu)) = t_r(\mu, Q^{-1}(\mu))$ is equivalent to

[33] To see this, note that $\Upsilon'(0.5) < 0$. Suppose $\mu^*_-(s) = \mu^*_+(s) = 0.5$, then $\Upsilon'(0.5) = 0$. Contradiction. Suppose $\mu^*_-(s) = \mu^*_+(s) < 0.5$, then $\Upsilon'(0.5) > 0$. Contradiction.

[34] Sorry to abuse the notation. We can verify that if $\mu \in [\underline{\mu}^*, 0, 5]$, $s^*(\mu)$ is the inverse function of $\mu^*(\cdot)$ after we prove this lemma.

[35] Note that equation (18), the exact expression of $t_r(\mu, s)$, is valid for all $\mu \in [0, 1]$.

find the solution to the system of equations below,

$$\begin{cases} \Pi_1(\mu, s) = 0, \\ V(\mu, D) = \mathbb{E}(v|\mu) - (v_h - D), \\ s = D, \end{cases}$$

which can be verified to have a unique non-negative solution $\mu = 0.5$. This suggests that $\Pi_1(\mu, D(\mu)) = \Pi_1(\mu, Q^{-1}(\mu)) = 0$ has a unique solution at 0.5. Moreover,

$$\frac{dt_r(\mu, D(\mu))}{d\mu} = \frac{\partial t_r(\mu, D)}{\partial \mu} + \frac{\partial t_r(\mu, D)}{\partial s}\frac{dD}{d\mu},$$

$$\frac{dt_r(\mu, s^*(\mu))}{d\mu} = \frac{\partial t_r(\mu, s^*)}{\partial \mu} + \frac{\partial t_r(\mu, s^*)}{\partial s}\frac{ds^*}{d\mu}.$$

Since $\frac{\partial t_r(\mu,s)}{\partial \mu}$ is independent of s, the first term of the two derivatives are the same. Besides, $\frac{dD}{d\mu} < 0$ and $\frac{ds^*}{d\mu} > 0$ if $\mu \leq 0.5$. Hence, the slope of $\bar{t}_r$ is smaller than $t_r(\mu, s^*(\mu))$. That is, if we reduce μ from 0.5, $t_r(\mu, s^*(\mu))$ decreases faster than $\bar{t}_r(\mu)$. Let $\underline{\mu}^*$ be the solution of $t_r(\mu, s^*(\mu)) = 0$. Obviously, $\underline{\mu}^* \in (\underline{\mu}, 0.5)$. To pin down $\underline{\mu}^*$, note that $t_r(\mu, s) = 0$ implies $s = q^{-1}(\mu)$. Thus, $\underline{\mu}^*$ is the solution that $\Pi_1(\mu, q^{-1}(\mu)) = 0$. Explicitly,

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

which also has a unique solution that $\underline{\mu}^* = \frac{k}{\lambda v_h} + (\frac{k}{\lambda v_h}(\frac{k}{\lambda v_h} + \frac{v_l}{v_h - v_l}))^{\frac{1}{2}}$.[36] Therefore, we pin down the set $[\underline{\mu}^*, 0.5]$ on which $t_r(\mu, s^*(\mu)) \in [0, \bar{t}_r(\mu)]$. $\square$

From this claim, we can see that $t_r(\mu, s^*(\mu)) > \bar{t}_r(\mu)$ if $\mu > 0.5$. Moreover, given that $\mu_+^*(s) > 0.5$, if there exists a local minimizer $\mu_+^*(s)$, it is larger than $Q(s)$. Therefore, $\Pi(\mu, s)$ is quasi-concave on $[q(s), Q(s)]$.

Denote $t_r^*(\mu) := t_r(\mu, s^*(\mu))$ for the domain $[\underline{\mu}^*, 0.5]$. Given the monotonicity of $s^*(\mu)$ when $\mu \leq 0.5$, we can conclude that if $s \in (q^{-1}(\underline{\mu}^*), Q^{-1}(0.5))$, the local maximizer $\mu_-^*(s) \in (q(s), Q(s))$ hence $\mu^*(s) = \mu_-^*(s)$ is the global maximizer. Besides, if $s \geq Q^{-1}(0.5)$, then $Q(s) \leq 0.5 \leq \mu_-^*(s)$, where the first inequality comes from $Q(s)$ being decreasing in s and the second inequality comes from $\mu_-^*(Q^{-1}(0.5)) = 0.5$. The inequality holds with equality only at $s = Q^{-1}(0.5)$. It is optimal to induce

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

a return belief $Q(s)$. If $s \leq q^{-1}(\underline{\mu}^*)$, $q(s) \geq \underline{\mu}^*$ and then $\Pi_1(q(s), s) \leq 0$.[37] Since $\Pi_1(\mu, s)$ is quasi-concave in $[q(s), Q(s)]$, thus if $\Pi_1(\mu, s) \leq 0$ at $q(s)$, $\Pi_1(\mu, s) \leq 0$ for all $[q(s), Q(s)]$. Still the inequality holds with equality only at $s = q^{-1}(\underline{\mu}^*)$. It is optimal to induce return belief $q(s)$. □

Theorem 1

Proof. Substituting the first-order condition (6) into the seller's expected revenue (7), we can simplify the latter expression to write:

$$\Pi(\mu, s^*(\mu)) = t_r^*(\mu) + \frac{\partial t_r(\mu, s^*(\mu))}{\partial \mu}(\mu_0 - \mu). \tag{19}$$

Taking the derivative w.r.t μ gives

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

Note that $\frac{\partial t_r(\mu, s^*(\mu))}{\partial \mu}$ is independent of s and we can verify $\frac{\partial^2 t_r^*}{\partial \mu^2} < 0$.[38] 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 is increasing in μ.

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

The proof of this claim can be found subsequent to this theorem. Recall Lemma 3, $\mu^*(s)$ is an optimal solution only for $s \in [q^{-1}(\underline{\mu}^*), Q^{-1}(0.5)]$. In particular, partial learning is optimal for $s \in (q^{-1}(\underline{\mu}^*), Q^{-1}(0.5))$; and for the boundaries, either full learning or no leaning is optimal. Consider the original problem ($\mathcal{P}$) and reimpose the two constraints: $q^{-1}(\mu_0) \leq s \leq Q^{-1}(\mu_0)$ and $\mu \leq \mu_0$, then $\mu^*(s)$ could be an optimal solution only if

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

which is equivalent to

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

[37] See footnote 41.

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

Figure 13 depicts this region. Note that when $\mu_0 \neq 0.5$, the upper boundary of $\mu^*(s)$, subject to the two constraints, is not 0.5. In particular, if $\mu_0 \leq 0.5$, then the optimal stopping belief $\mu \leq \mu_0$ (see Figure 13 (a)); if $\mu_0 > 0.5$, then the optimal stopping belief $\mu \leq \mu^*(Q^{-1}(\mu_0))$ (see Figure 13 (b)), The lower boundary $\underline{\mu}^*$ can always be achieved when $\mu_0 \in [\underline{\mu}^*, Q(q^{-1}(\underline{\mu}^*))]$.

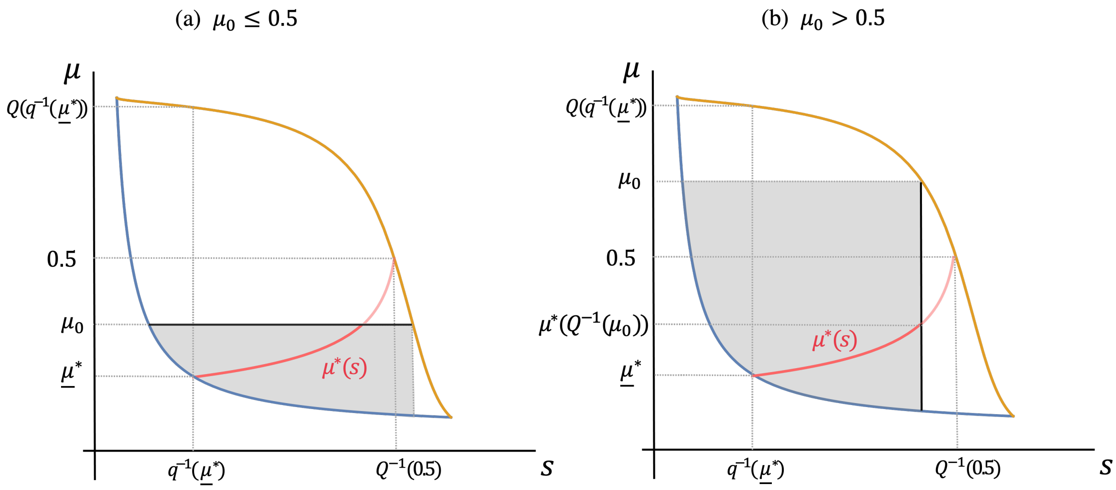


Figure 13: Partial learning as optimal interior solution

We distinguish two cases. First, $\phi'(\underline{\mu}^*) \geq -1$ implies that $\Pi(\mu, s^*(\mu))$ is quasi-convex in μ. Second, $\phi'(\underline{\mu}^*) < -1$ implies 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.

Case one: $\phi'(\underline{\mu}^*) \geq -1$. This is true in most scenarios. 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, as depicted in Figure 14, where the black lines represent the contour lines of $\mu_0 - \mu$ for different μ_0.

- If $\mu_0 < \Phi(\underline{\mu}^*)$, then $\Pi(\mu, s^*(\mu))$ is increasing in μ. This implies that inducing the upper boundary of $\mu^*(s)$, subject to the two constraints, is optimal, which further implies the optimality of Learning Deterrence. To see this, when $\mu_0 \leq 0.5$, the optimal return belief is μ_0 and inducing no learning via Stochastic Return is strictly dominated by Learning Deterrence. When $\mu_0 > 0.5$, the optimal return belief is $\mu^*(Q^{-1}(\mu_0)) < 0.5 < \mu_0$. However, we

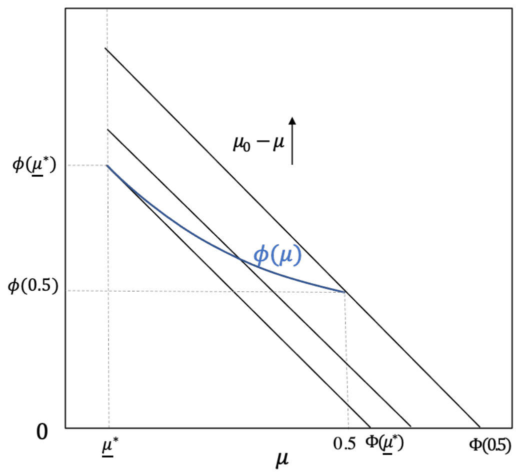


Figure 14: Case one: seller's expected revenue is quasi-convex

can show that Learning Deterrence is better than Stochastic Return that induces stopping at $\mu^*(Q^{-1}(\mu_0))$. Specifically, 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 $t^D(0.5) = v_h - Q^{-1}(0.5)$ and the return transfer $t_r(0.5, Q^{-1}(0.5))$, which is smaller than $t^D(0.5)$ and since $t^D(\cdot)$ is increasing, $t^D(\mu_0) > t^D(0.5)$, meaning that Learning Deterrence generates strictly higher profit.[39]

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

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

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 15 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 that if $\mu_0 \in [\Phi(r), \Phi(\underline{\mu}^*)]$,

$$\Pi(r_1(\mu_0), s^*(r_1(\mu_0))) < t^D(\mu_0).$$

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

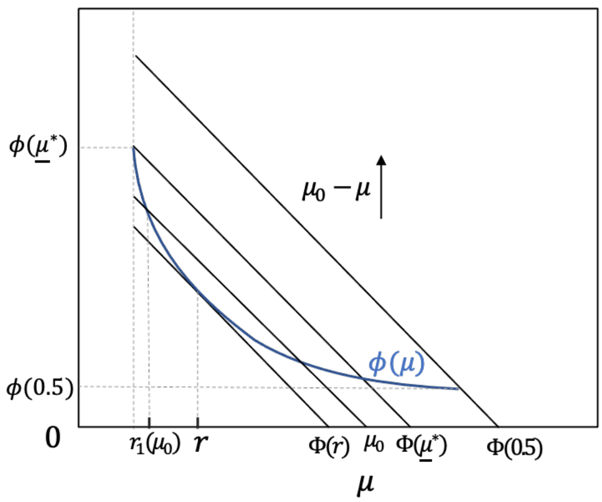


Figure 15: Case two: seller's expected revenue is not quasi-convex

With slight abuse of notation, 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 Π increasing in μ_0. The second inequality is due to $\underline{\mu}^* = r_1(\Phi(\underline{\mu}^*))$, which is the local maximizer of Π when $\mu_0 = \Phi(\underline{\mu}^*)$. Recall equation (19) and plug in the expression of $\underline{\mu}^*$,

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

It is obvious that $s < v_h$ whenever learning is feasible. Thus,

$$\mathbb{E}(v|\frac{k}{\lambda v_h}) < \mathbb{E}(v|\underline{\mu}) = t^D(\underline{\mu}) < t^D(\mu_0),$$

where the equality and the second inequality come from Proposition 2. □

Claim 3

Proof. Denote $w(\mu) := \mathbb{E}[v|q(s^*(\mu))]$, then $\phi(\mu) = \frac{1-\mu}{v_h} w(\mu)$. Note that $w(\mu)$ is decreasing in μ, as $q(s)$ decreases in s and $s^*(\mu)$ increases in μ. Besides, we can verify that $s^*(\mu)$ is concave for $\mu \in [\underline{\mu}^*, 0.5]$,[40] hence $w(\mu)$ is convex.

[40]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 $\frac{d^2 s^*}{d\mu^2} < 0$.

Note that

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

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 in μ. That is, $\phi(\mu)$ is convex.

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

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

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 $\phi'(0.5)|_{v_l \to v_h} = -1$.[41] □

Theorem 2

Proof. First, we want to show that F is either an empty set or a closed interval. Note that $t^D(\underline{\mu}) > \Pi^F(\underline{\mu})$ and $t^D(\overline{\mu}) > \Pi^F(\overline{\mu})$. Hence, it is equivalent to show $\Pi^F(\mu_0)$ crosses $t^D(\mu_0)$ at most twice. Let $\Pi^F(\mu) = t^D(\mu) = v_h - D(\mu)$, then $D(\mu) = v_h - \Pi^F(\mu)$. To simplify the exposition, let $\theta(\mu) := \frac{\gamma}{(v_h - \Pi^F(\mu))}$. Recall equation (12) and $\tilde{\mu}(\mu) = \frac{\gamma}{D(\mu)}$. Then we want to show $g(\mu) - f(\theta(\mu))$ has at most two roots when $\mu \in [\underline{\mu}, \overline{\mu}]$. To verify this,

$$g' - f'\theta' = \frac{1}{1-\mu}\left(\frac{1}{1-\mu} + \frac{1}{\mu}\right) + \left(\frac{1}{\sqrt{\mu}} - \frac{r}{\sqrt{1-\mu}}\right)\frac{(\sqrt{\mu} + r\sqrt{1-\mu})^3}{(\sqrt{\mu} + r\sqrt{1-\mu})^2 - 1},$$

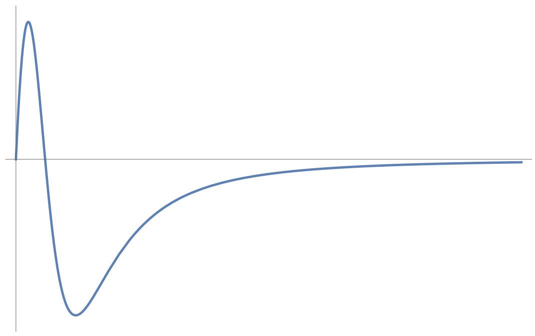
where $r = \sqrt{v_h/\gamma - 1} > \sqrt{3}$ given the assumption that $v_h > 4\gamma + v_l$. Let $x = \sqrt{\frac{\mu}{1-\mu}} \in (0, \infty)$, which is a monotone transformation of μ. Rearranging $g' - f'\theta' = 0$, we have

$$m(x) := \frac{x(x+r)^3(1-xr)}{(1+x^2)^3(-1+2xr+r^2)} = -1,$$

where $m(x)$ is a rational function. The degree of the numerator is smaller than that of the denominator, thus it has a horizontal asymptote $m = 0$. Note that the denominator is positive due to $\theta(\mu) \in [0, 1]$, hence it does not have a vertical

[41] 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 $\frac{d\phi'(0.5)}{dv_l} = \frac{1}{v_h^2}[(q^*(0.5) - 1)(v_h + 4(v_h - 2v_l)q^*(0.5)^2) - (v_h - v_l)[v_h + 4v_l(2 - 3q^*(0.5))q^*(0.5)]\frac{dq^*(0.5)}{dv_l}] < 0$.

asymptote. Meanwhile $\lim_{x\to 0} m(x) = 0$, $\lim_{\mu\to\infty} m(x) = 0$, $m(x=1) < 0$, and $m(x) = 0$ has a unique root $x = 1/r < 1$. Therefore, the graph of $m(x)$ is the following.



Then, $m(x) = -1$ has at most two roots. That is, if $\mu \in [\underline{\mu}, \overline{\mu}]$, $g'(\mu) - f'\theta'(\mu)$ has at most two roots and $g'(\mu) - f'\theta'(\mu) < 0$ when μ is between the two roots. Given that $g(\mu) - f(\theta(\mu))$ is strictly positive at $\underline{\mu}$ and $\overline{\mu}$, we can verify $g(\mu) - f(\theta(\mu))$ has at most two roots. The existence of γ^* is implied by Proposition 4, and the limit result in Proposition 5 implies that the left endpoint of F is larger than $\frac{v_l}{v_h}$. The exact form of optimal refund mechanism is an immediate result of Corollary 1 and Theorem 1. □

Proposition 4

Proof. Recall that $\Pi(q(s), s) = \frac{\mu_0 - \gamma/s}{1-\gamma/s}(v_h - s)$. By the 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 $\tilde{\mu}(\mu) = q(D(\mu))$. Taking the 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 closed 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,

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

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

Proposition 5

Proof. First we calculate the limit of $t^D(\mu_0)$ when $\gamma\to 0$. Plugging $D(\mu_0)=\frac{\gamma}{\bar{\mu}(\mu_0)}$ into equation (12) and multiplying by γ gives:

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

If $\gamma\to 0$ and μ_0 does not converge to 0 or 1, the above equation converges to $v_h-D(\mu_0)=v_l$.[42] 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 $\mu_0=\frac{v_l}{v_h}$.

Second, 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 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 $\frac{v_l}{v_h}\le\mu_0\ll 1$, $\lim_{\gamma\to 0}\Pi^F(\mu_0)>\lim_{\gamma\to 0}t^D(\mu_0)$. Then, when $\mu_0\ge\frac{v_l}{v_h}$, the seller's expected revenue converges to $\mu_0 v_h$. □

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

Proposition 6

Proof. For any fixed s, the lower bound of the buyer's ex-ante surplus is still $V^0(\mu_0, s; \lambda_B)$ as the buyer can always treat the mechanism as if a return is prohibited. Note that for a mechanism $\{t_b, (x_r, t_r + t_u)\}$, the buyer's value function for post-transaction learning is $V_P(\mu, s + t_u; \lambda_P) - t_u$. This is obtained by simple normalization.

Now we show there is a profitable deviation if the seller encourages post-transaction learning and provides the buyer with a surplus strictly larger than $V^0(\mu_0, s; \lambda_B)$. Note that when the mechanism is $\{t_b, (x_r, t_r + t_u)\}$, the seller with reservation value u obtains the following return revenue by inducing the buyer to stop learning and request a return at belief μ,

$$t_r + t_u = \mathbb{E}(v|\mu)\frac{V'_P(\mu, s + t_u; \lambda_P)}{v_h - v_l} - (V_P(\mu, s + t_u; \lambda_P) - t_u), \tag{20}$$

where V'_P represents the partial derivative w.r.t to μ. From the (ODE) for post-transaction learning with a normalized surplus $s + t_u$, we obtain

$$V_P(\mu, s + t_u; \lambda_P) = s + t_u - \frac{k}{\mu\lambda_P} - (1 - \mu)V'_P(\mu, s + t_u; \lambda_P).$$

Thus, equation (20) can be further reduced to

$$t_r + t_u = \frac{v_h}{v_h - v_l}V'_P - s + \frac{k}{\mu\lambda_P}$$

Note that if lowering V_P, V'_P is increasing by the above differential equation. Besides, the return revenue is increasing in V'_P. Hence, if $V_P(\mu, s+t_u; \lambda_P) - t_u > V^0(\mu, s; \lambda_B)$, the seller can gain larger expected revenue by raising the cancellation fee t_u while let the buyer preserve the same stopping belief. Thus, the optimality holds when equation (8) holds.

Next, it is very easy to verify that for fixed s, the seller's return payoff $t_r + t_u$ is always larger under post-transaction learning than inducing the the buyer to stop at the same belief but under pre-transaction learning. To induce the same stopping belief $\mu < \mu_0$ while restricting $V(\mu_0, s + t_u; \lambda_P) - t_u = V(\mu_0, s; \lambda_B)$, we can verify $V'(\mu, s + t_u; \lambda_P) > V'(\mu, s; \lambda_B)$. Therefore, the allocation rate at return is larger and the buyer's continuation value is lower, implying the seller obtains larger return revenue by inducing the same stopping belief with post-transaction learning.

□

Proposition 7

Proof. From now on, we write $t_u(s,\mu_0)$ as the solution to equation (8). Denote $q_B^{-1}(\mu_0) := \frac{k}{\lambda_B \mu_0}$ which is the inverse of quitting belief for pre-transaction learning. Similarly, denote $q_P^{-1}(\mu_0) := \frac{k}{\lambda_P \mu_0}$ and $Q_B^{-1}(\mu_0) = Q^{-1}(\mu_0;\lambda_b)$. Notice that (1) when $s \in [q_P^{-1}(\mu_0), q_B^{-1}(\mu_0))$, $V(\mu_0, s+t_u;\lambda_P) - t_u = V^0(\mu_0, s;\lambda_B) \equiv 0$. And (2) if $s > q_B^{-1}(\mu_0)$, $V(\mu_0, s+t_u;\lambda_P) - t_u = V^0(\mu_0, s;\lambda_B) > 0$. Therefore the return revenue $t_r(\mu,s)+t_u(s,\mu_0)$ follows different expression for the above two cases. We discuss them separately.

Case one: $s \in [q_P^{-1}(\mu_0), q_B^{-1}(\mu_0))$. Substitute equation (8) into equation (20), we obtain:

$$t_r(\mu,s)+t_u(s,\mu_0) = \frac{1}{\lambda_P \mu (v_h - v_l)}\Big[\frac{\mu(k-\lambda_P s\mu_0)v_h}{\mu_0 - 1} - kv_l + \lambda_P s\mu v_l \\ + k\mu v_h\left(\log\left[\frac{\mu}{1-\mu}\right] - \log\left[\frac{\mu_0}{1-\mu_0}\right]\right)\Big].$$

We can verify that

$$\lim_{\lambda_P\to\infty} t_r(\mu,s)+t_u(s,\mu_0) = \left(\frac{v_h}{(1-\mu_0)(v_h-v_l)} - 1\right)s$$

Thus, the seller's expected revenue equals,

$$\lim_{\lambda_P\to\infty} \mu_0(v_h - s) + (1-\mu_0)[t_r(\mu,s)+t_u(s,\mu_0)] = \mu_0 v_h + \frac{v_l s}{v_h - v_l}$$

which is increasing in s. Therefore the seller optimally sets an $s = q_B^{-1}(\mu_0)$ in this case.

Case two: $s \in [q_B^{-1}(\mu_0)), Q_B^{-1}(\mu_0))$. Substitute equation (8) into equation (20), we obtain:

$$t_r(\mu,s)+t_u(s,\mu_0) = -\frac{kv_h \log\left[\frac{k}{\lambda_B s - k}\right]}{\lambda_B(v_h - v_l)} + \frac{-kv_l + \lambda_P s\mu v_l + k\mu v_h \log\left[\frac{\mu}{1-\mu}\right]}{\lambda_P \mu (v_h - v_l)} \\ + \frac{k(\lambda_P - \lambda_B)v_h(1 + (1-\mu_0)\log\left[\frac{\mu_0}{1-\mu_0}\right])}{\lambda_B\lambda_P(1-\mu_0)(v_h - v_l)}.$$

We can verify that

$$\lim_{\lambda_P\to\infty} t_r(\mu,s)+t_u(s,\mu_0) = \frac{kv_h + \lambda_B s(1-\mu_0)v_l + k(1-\mu_0)v_h\left(\log[\frac{\mu_0}{1-\mu_0}] - \log[\frac{k}{\lambda_B s - k}]\right)}{\lambda_B(1-\mu_0)(v_h - v_l)}$$

Take the first order derivative of $\lim_{\lambda_P \to \infty} \mu_0(v_h - s) + (1-\mu_0)[t_r(\mu, s) + t_u(s, \mu_0)]$ w.r.t s, we obtain,

$$\frac{\lambda_B(1-\mu_0)v_h s}{(\lambda_B s - k)(v_h - v_l)} - 1,$$

which is decreasing in s. Therefore, the seller's revenue is increasing in s if $s \leq \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$, otherwise, it is decreasing in s if $s > \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$.

(1) If $\mu_0 \leq v_l/v_h$, then $s > 0 > \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$ and the seller's revenue is always decreasing in s, which renders $s = q_B^{-1}(\mu_0)$ the locally optimal solution. However, we can verify that the locally optimal solution is dominated by setting $s = Q_B^{-1}(\mu_0)$ to deter learning. (2) If $\mu_0 \in (v_l/v_h, \overline{\mu}]$, then $\frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)} > q_B^{-1}(\mu_0)$. We can verify that there exists μ_0' and μ_0'' such that $v_l/v_h < \mu_0' < \mu_0'' < \overline{\mu}$, and if $\mu_0 \in [\mu_0', \mu_0'']$, then $q_B^{-1}(\mu_0) \leq \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)} \leq Q_B^{-1}(\mu_0)$, rendering the optimal solution $s = \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$. Otherwise, if $\mu_0 \notin [\mu_0', \mu_0'']$, the optimal solution is $s = Q_B^{-1}(\mu_0)$. Given this, suppose the optimal $s = \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$. Then the optimal price when $\lambda_P \to \infty$ is

$$v_h - \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}.$$

The optimal cancellation fee when $\lambda_P \to \infty$ is

$$\frac{k}{\lambda_B(1-\mu_0)} \left\{ 1 + (1-\mu_0) \log \left[\frac{\mu_0 v_h}{\mu_0 v_h - v_l} \right] \right\}.$$

The optimal return transfer when $\lambda_P \to \infty$ is

$$\frac{k v_l \left(v_h - v_l - (1-\mu_0) v_l - (1-\mu_0)(\mu_0 v_h - v_l) \log \left[\frac{\mu_0 v_h}{\mu_0 v_h - v_l} \right] \right)}{\lambda_B (1-\mu_0)(v_h - v_l)(\mu_0 v_h - v_l)}.$$

The optimal allocation rate at return when $\lambda_P \to \infty$ is

$$\frac{k \left(v_h - v_l - (1-\mu_0) v_l - (1-\mu_0)(\mu_0 v_h - v_l) \log \left[\frac{\mu_0 v_h}{\mu_0 v_h - v_l} \right] \right)}{\lambda_B (1-\mu_0)(v_h - v_l)(\mu_0 v_h - v_l)}.$$

□

References

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.

Board, Simon, "Selling Options," *Journal of Economic Theory*, 2007, *136*, 324–340.

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

Che, Yeon-Koo, "Customer Return Policies for Experience Goods," *The Journal of Industrial Economics*, 1996, *44* (1), 17–24.

Courty, Pascal and Hao Li, "Sequential Screening," *The Review of Economic Studies*, 10 2000, *67* (4), 697–717.

Daley, Brendan, Thomas Geelen, and Brett Green, "Due Diligence," *Working paper*, 2021.

Hinnosaar, Toomas and Keiichi Kawai, "Robust pricing with refunds," *The RAND Journal of Economics*, 2020, *51* (4), 1014–1036.

Inderst, Roman and Gilad Tirosh, "Refunds and returns in a vertically differentiated industry," *International Journal of Industrial Organization*, 2015, *38*, 44–51.

_ **and Marco Ottaviani**, "Sales Talk, Cancellation Terms and the Role of Consumer Protection," *The Review of Economic Studies*, 2013, *80* (3 (284)), 1002–1026.

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 and Sven Rady, "Breakdowns," *Theoretical Economics*, 2015, *10*, 175–202.

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

Krähmer, Daniel and Roland Strausz, "Optimal Sales Contracts with Withdrawal Rights," *The Review of Economic Studies*, 2015, *82* (2 (291)), 762–790.

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.

Matthews, Steven A. and Nicola Persico, "Information Acquisition and Refunds for Returns," *(Unpublished Manuscript, University of Pennsylvania)*, 2007.

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*, 2020.

Pomatto, Luciano, Philipp Strack, and Omer Tamuz, "The cost of 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.

Shieh, Shiou, "Price and Money-Back Guarantees as Signals of Product Quality," *Journal of Economics & Management Strategy*, 1996, *5* (3), 361–377.

Wei, Dong and Brett Green, "(Reverse) Price Discrimination with Information Design," *working paper*, 2020.