ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 214

Optimal Refund Mechanism

Qianjun Lyu

December 2022

www.econtribute.de

UNIVERSITÄT BONN

Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under
Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

Optimal Refund Mechanism

Qianjun Lyu*

University of Bonn

December 21, 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 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 good news and bad news framework.

Keywords: buyer learning, refund contract, information design, implementable mechanism.

*Institute for Microeconomics, University of Bonn. I would like to thank Wing Suen for insightful and continual discussions. For valuable suggestions, I would like to thank Pak Hung Au, Francesc Dilmé, Bard Harstad, Wei He, Jin Li, Yunan Li, Xianwen Shi, Dezsö Szalay, Balázs Szentes, Jianrong Tian, Dong Wei, Yimeng Zhang and Chen Zhao. I also thank seminar and conference participants at University of Bonn, LSE, University of Nottingham and Econometric Society meetings for helpful comments. I acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1– 390838866.

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.[1] 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

[1] In terms of implementing such refund mechanisms, the seller who receives multiple return forms can request only a proportion of buyers to return the product.

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 over time both before and after purchase. One can interpret the before-purchase learning as the product exploration phase and the after-purchase learning as the evaluation phase. Specifically, at each instant of time, based on the information already collected, the buyer can decide whether to keep acquiring information or stop learning and make a decision. In the main model, we study the "no news is bad news" benchmark, i.e., by exerting costly effort, good news arrives at some Poisson rate if buyer's true valuation is high, otherwise no news arrives if his true valuation is low. We assume the learning rate after purchase is weakly higher than the learning rate before purchase. Moreover, letting the after-purchase learning rate go to infinity, we can accommodate the case where the buyer immediately learns his true valuation 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. With the dynamic feature of the buyer's learning process, the seller can implement more flexible information acquisition as the buyer's decision for stopping learning is endogenously determined by the refund mechanism. Essentially, a refund mechanism is an option contract as it offers the buyer two options: either to consume the item and obtain the consumption utility, or to return it and obtain the return payoff. The buyer optimally decides when to stop learning while comparing the continuation value from learning and the maximum payoff from consumption and return.

With the exponential-bandit framework (Keller, Rady and Cripps (2005)), the buyer's non-degenerate posterior belief upon stopping determines the total amount of information the buyer acquires. Yet, the buyer's stopping belief is unobserved to the seller. We study *implementable* refund mechanism—a mapping from the buyer's posterior belief into an allocation rule—such that (1) the buyer truthfully reports his private posterior belief at stopping; and (2) the buyer optimally stops learning at the posterior belief reported. Given this, we can transform a mechanism design problem into an information design problem. In particular, we first characterize the set of posterior beliefs that are implementable across different prices. Next, we specifying the refund mechanism that can implement each posterior belief. Finally, we let the seller maximizes over the set of implementable posterior beliefs.

For each price, the set of inducible posterior beliefs can be segmented into three groups, namely, 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 put it differently, the seller prevents the buyer from private learning. To achieve this, the seller sets a sufficiently low non-refundable price just to make the buyer indifferent between consuming the item and continuing to learn. In the last case, the first-best is achieved.

Our main result in the benchmark model suggests that the optimal stopping belief that the seller aims to implement is located at the boundaries of the implementable set of beliefs, given that the price is also optimally chosen. That is, inducing partial learning is sub-optimal, further implying the optimality of a deterministic mechanism:[2] if the seller allows a return, she requires the buyer to return the product with probability one; otherwise, she does not allow a return. 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. 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 issuing a 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 learning or encourages full learning via a 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. By decreasing the price just a little, the seller can deter learning and induce immediate consumption, thereby to capture a large fraction of the first-best allocation surplus. How-

[2]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.

ever, if the buyer's prior belief becomes more uncertain, then information values a lot, which makes deterring learning less profitable as it requires a significant price reduction for the compensation of buyer's opportunity information rent. On the other hand, encouraging learning becomes more appealing as it avoids such a compensation. In other words, the seller can significantly raise 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 belief is more optimistic. As a result, the seller optimally allows a free return if the buyer's prior belief is less extreme but relatively more optimistic. The tension between welfare maximization and profit maximization is mainly driven by the buyer's option value from information acquisition.

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.

If the learning rate after purchase becomes higher, then a cancellation fee (partial refund) is involved if the seller allows a return. We characterize the optimal refund mechanism when the learning rate after purchase converges to infinity so that the buyer can almost learn his true valuation immediately. In the case that $v_l = 0$, the optimal refund mechanism either deters learning with a sufficiently low price or allows a return but charges a cancellation fee. Both mechanisms are common in fashion online platform. For example, the former one corresponds to "Final sales" and the latter one corresponds to "Return with a fixed fee".

In another extension where we look at the "no news is good news" case. The optimal refund mechanism turns out to have the same structure as in the benchmark model. However, with this learning technology, if the seller allows a free return, she escalates price discrimination to the extreme in the sense that she fully extracts the buyer's ex-ante surplus. In other words, the buyer receives *zero* surplus if the seller optimally allows a free return.

Related literature. Our paper relates to the sequential screening literature. Courty and Li (2000) study how refund contract price discriminates buyers who have asymmetric ex-ante imperfect private information but can observe the true valuation after contracting. In contrast, we consider symmetric ex-ante information and study how refund contract elicits buyer's ex-post private information. Krähmer and Strausz (2015) impose buyer's ex-post participation constraints in the standard sequential screening model, while in our paper, the availability of ex-post participation is the seller's choice.

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, the Markov nature of our learning technology allows us to analyze how the seller's optimal mechanism varies with the buyer's prior belief, which cannot otherwise be accommodated in the flexible information cost framework.[3] In terms of sequential buyer learning,[4] Lang (2019) and Pease (2020) investigate the seller's pricing policy when the buyer can acquire information over time.

From a robustness perspective, 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. Roesler and Szentes (2017) studies the buyer's optimal information design anticipating its impact on the seller's pricing decision. Ravid, Roesler and Szentes (2019) consider the same scenario but let both seller and buyer move simultaneously so as to discuss the equilibrium outcomes. 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. They characterize seller's best guaranteed profit.

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 flexibility in manipulating the

[3]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).

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

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 the asset at a contingent fee or to 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.

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. Let $\mu_0 = \Pr(v_h)$ be the common prior belief that the product valuation is high. We use μ to represent the buyer's posterior belief and sometimes call this the buyer's type. A type-μ buyer's expected product valuation is $\mathbb{E}[v|\mu] := \mu v_h + (1-\mu)v_l$. The buyer's type evolves over time; 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.[5]

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 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 probability 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. Under a typical refund mechanism $\{t_b, (x_r, t_r)\}$, 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.

[5] This is not crucial to the analysis as the seller obtains an upfront payment even with a free return. Nevertheless, we assume it away because the time between purchase and return is usually not very long.

Given this notation, under a No Return mechanism, $x_r = 1$, as the buyer cannot return the item and therefore always has to keep it. Conversely, under a Free Return mechanism, $x_r = 0$ 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 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. In particular, a type-μ buyer obtains the consumption utility $\mathbb{E}[v|\mu] - t_b$ if he purchases the item without requesting a return, or the return utility $\mathbb{E}[v|\mu]x_r - t_r$ 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:

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

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. If the seller does not allow a return, the buyer acquires information before purchase. Denote $V^0(\mu(\tau); t_b, 1, t_b)$ as the buyer's value function given a No Return mechanism $\{t_b, (1, t_b)\}$. It is determined by the Bellman equation below,

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

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 $\mu(\tau + d\tau)$. The solution to (2) determines the buyer's learning strategy under No Return mechanisms.

Suppose that the seller allows a return with policy (x_r, t_r) and the buyer purchases the item first and acquires information afterwards. Then the buyer's value function for purchase $V_P(\mu(\tau); t_b, x_r, t_r)$ is determined by the Bellman equation,

$$\begin{aligned} V_P(\mu(\tau); t_b, x_r, t_r) = \max\{\mathbb{E}[v|\mu(\tau)] - t_b, \mathbb{E}[v|\mu(\tau)]x_r - t_r, \\ -kd\tau + \mu(\tau)\lambda d\tau(v_h - s) + (1-\mu(\tau)\lambda d\tau)V_P(\mu(\tau+d\tau); t_b, x_r, t_r)\}. \end{aligned} \tag{3}$$

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. The solution to (3) determine the buyer's learning strategy after he purchases the item.

2.1 Implementable mechanism

Since the seller cannot observe the buyer's posterior belief when the buyer stops learning, then for a refund mechanism to implement some particular stopping belief, we require the buyer to be willing to stop learning at the belief he reports to the seller. Specifically, the return policy, $x_r(\mu) : [0, \mu_0] \to [0, 1]$ and $t_r(\mu) : [0, \mu_0] \to [0, t_b]$, assigns the allocation rule conditional on the buyer's report on his private stopping belief. On the other hand, the price essentially describes an allocation rule for a report of belief 1 (after the arrival of good news). To guarantee the optimality of a refund mechanism with non-trivial return policies,

$$V_P(\mu_0; t_b, x_r(\mu), t_r(\mu)) \geq V^0(\mu_0; t_b, 1, t_b). \tag{IR}$$

Because if it fails, the buyer acquires all information before purchase and no return request can be realized. In other words, buyer's value from learning under a No Return mechanism imposes a lower bound on his expected trading surplus.

Call a refund mechanism $\{t_b, (x_r(\mu), t_r(\mu))\}$ *implementable in terms of encouraging learning* if the following incentive constraints hold.

$$V_P(\mu_0; t_b, x_r(\mu), t_r(\mu)) > \max\{\mathbb{E}[v|\mu_0] - t_b,\ \mathbb{E}[v|\mu_0]x_r(\mu) - t_r(\mu)\}. \tag{IM-L1}$$

$$\inf\left\{\mu' \in [0, \mu_0] : V_P(\mu'; t_b, x_r(\mu), t_r(\mu)) > \mathbb{E}[v|\mu']x_r(\mu) - t_r(\mu)\right\} = \mu. \tag{IM-L2}$$

$$\mathbb{E}[v|\mu]x_r(\mu) - t_r(\mu) \geq \mathbb{E}[v|\mu] - t_b \text{ and } v_h - t_b \geq v_h x_r(\mu) - t_r(\mu). \tag{IC}$$

Constraint (IM-L1) means that the continuation value for learning at the prior belief is higher than the maximum payoff from consumption and return, therefore the buyer is willing to learn at the prior belief. Constraint (IM-L2) implies that the buyer is willing to stop learning at the belief he reports to the seller. Notice that V_P is implicitly determined given buyer's optimal learning strategy. We will elaborate more on how we derive it later. At this stage, it is useful to notice that if learning is optimal at the prior belief, then there exists an intermediate interval of beliefs such that learning is optimal. Constraints (IC) is the incentive constraints for truth-telling such that a type-μ buyer prefers a return rather than consumption and a type-1 buyer prefers consumption rather than a return. Interestingly, (IM-L1) and (IM-L2) imply the incentive compatibility for truth-telling (IC) because otherwise acquiring information is unnecessary.

Conversely, call a refund mechanism $\{t_b, (x_r(\mu), t_r(\mu))\}$ *implementable in terms of deterring learning* if the following incentive constraints hold, i.e., the buyer weakly prefers to consume the item rather than continuing to learn.

$$V_P(\mu_0; t_b, x_r(\mu), t_r(\mu)) = \mathbb{E}[v|\mu_0] - t_b. \tag{IM-D}$$

The seller maximizes the expected revenue subject to the implementable constraints for encouraging learning or deterring learning. The left program considers encouraging learning, while the right program considers deterring learning.

$$\sup_{t_b, x_r(\cdot), t_r(\cdot), \mu} \quad \frac{\mu_0 - \mu}{1-\mu} t_b + \frac{1-\mu_0}{1-\mu} t_r(\mu) \qquad \text{and} \qquad \sup_{t_b, x_r(\cdot), t_r(\cdot), \mu} \quad t_b$$
$$s.t. \quad \text{(IR), (IM-L1), (IM-L2)} \qquad\qquad s.t. \quad \text{(IR), (IM-D)}$$

While encouraging the buyer to learn, seller's expected revenue is a weighted average between the price and the return transfer, where $\frac{\mu_0-\mu}{1-\mu}$ is the ex-ante expected probability that the buyer obtains good news before his belief falls below μ, and $\frac{1-\mu_0}{1-\mu}$ is the expected probability for the complementary event. Conversely, while preventing type-μ_0 buyer from private learning, the seller's revenue equals the price. Note that the two objective functions are not equal when $\mu = \mu_0$.

Lemma 1. *Constraint (IR) binds.*

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, suppose the seller wants to encourage learning and designs a benevolent return policy that provides the buyer with a continuation value strictly larger than $V^0(\mu_0; t_b, 1, t_b)$. Suppose the seller wants to implement a stopping belief μ. Then she can increase the return transfer $t_r(\mu)$ and adjust the return allocation probability $x_r(\mu)$ properly such that the buyer's stopping belief remains the same. It implies a profitable deviation. If the seller deters learning, then this Lemma holds trivially.

3 A No Return Benchmark

Given Lemma 1, the buyer's learning strategy under a No Return mechanism serves as a building block for subsequent results. Without a return, the seller's only choice is the price. It endogenously determines the net consumer surplus $s \equiv v_h - t_b$ upon the arrival of good news (e.g., endogenous "prize" for breakthrough), which further determines the value of experimentation. For the sake of illustration and comparing results with the experimentation literature, we abuse the notation a little and use $V^0(\mu(\tau), s)$ to represent the value function under no return, and rewrite (2) as below,

$$\begin{aligned} V^0(\mu(\tau), s) = \max\{ \, 0 \, , \, \mathbb{E}(v|\mu(\tau)) - (v_h - s), \\ -kd\tau + \mu(\tau)\lambda d\tau s + (1-\mu(\tau)\lambda d\tau) V^0(\mu(\tau + d\tau), s)\}. \end{aligned} \tag{4}$$

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 $q(s)\leq Q(s)$, that determine the buyer's optimal learning strategy. Specifically, 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,[6] 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:

$$Q(s)=\{\mu : V(\mu,s)=\mathbb{E}[v|\mu]-(v_h-s)\}. \tag{5}$$

With slight abuse of notation, in equation (5) 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 to accept the price. Specifically,

$$\mathbb{E}(v|q(s))-(v_h-s)\leq 0. \tag{6}$$

If this inequality fails, no learning can be induced because learning has no value when it does not affect the purchase decision. Assumption 1 ensures that there exists a price in $[v_l,v_h]$ such that information acquisition is valuable for some prior belief. Without it, the buyer does not learn regardless of the prior belief. Thus, the seller sets a non-refundable price equal to buyer's ex-ante expected valuation $\mathbb{E}[v|\mu_0]$ to capture the entire allocation surplus.

Assumption 1. $(v_h-v_l)\lambda>4k$.

To avoid trivial result, we make this assumption throughout the paper. It implies there exist two distinct roots $\underline{s}<\bar{s}$ such that if $s\in[\underline{s},\bar{s}]$, acquiring information is optimal for some prior belief.

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

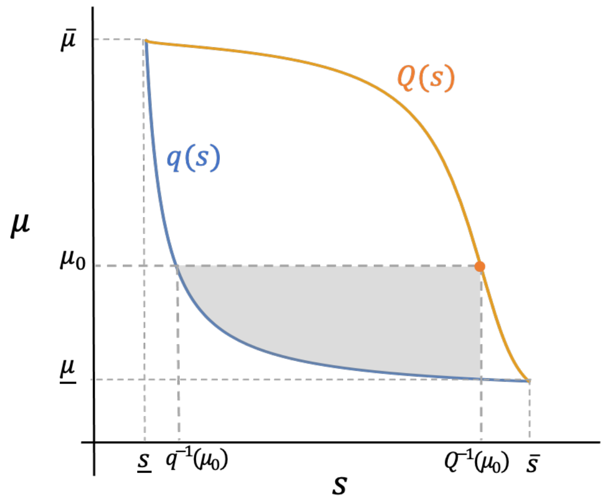


Figure 1: Inducible learning outcomes

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

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

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 $s \in [\underline{s}, \bar{s}]$, 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 belief $q(s)$. Figure 1 plots the quitting belief $q(s)$ and trial belief $Q(s)$ against s, which gives a complete characterization of buyer's learning strategy across different s. Both beliefs are decreasing in s and coincide at the two boundaries.[7] Denote $\underline{\mu} = q(\bar{s}) = Q(\bar{s})$ and $\bar{\mu} = q(\underline{s}) = Q(\underline{s})$.

Given Proposition 1, if $\mu_0 \notin [\underline{\mu}, \bar{\mu}]$, 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 buyer's ex-ante expected valuation, which

[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. Therefore, the belief interval that the buyer finds consumption optimal becomes larger. $q(s) = Q(s)$ at the boundaries is implied by inequality (6).

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 decision about the optimal amount of information she wants the buyer to acquire.

To study this, we characterize the set of implementable buyer stopping beliefs across s, and we call it the inducible learning outcomes. The shaded area in Figure 1 is the 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 implements the stopping belief $\mu \in [q(s), Q(s)]$ in the sense that the buyer is willing to stop learning at the belief he reports to the seller. We require $s \in [q^{-1}(\mu_0), Q^{-1}(\mu_0)]$ so that the type-μ_0 buyer (weakly) prefers to acquire information at the prior belief, implied by $V^0(\mu_0, s) \geq \max\{0, \mathbb{E}[v|\mu_0] - (v_h - s)\}$. Meanwhile, the stopping belief μ is bounded by μ_0 since no news is bad news.

In Section 4, we study the refund mechanisms that implement the boundaries of the shaded area in Figure 1, while in Section 5, we study the refund mechanisms that implement its interior part. Eventually, the seller maximizes her expected revenue over the entire shaded area.

4 Learning Deterrence and Free Return

Now we characterize the refund mechanisms that can implement the boundaries of the inducible learning outcomes. We are particularly interested in (1) the intersection of the shaded area and the orange curve $Q(s)$ (the orange dot in Figure 1); and (2) the intersection of the shaded area and the blue curve $q(s)$ in Figure 1.[8]

Learning Deterrence is a No Return mechanism with price $t^D(\mu_0) := v_h - Q^{-1}(\mu_0)$.[9] It is the solution to the seller's optimization for deterring learning. Under this mechanism, the type-μ_0 buyer is just indifferent between acquiring information and consuming

[8]We can easily verify other boundary points to be sup-optimal.

[9]Note that any non-refundable prices strictly lower than $t^D(\mu_0)$ can induce immediate consumption, but the seller then has an incentive to increase the price.

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 for the buyer's compensation until the value of information becomes non-positive.[10] 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, Learning Deterrence induces efficient allocation and therefore achieves the first best solution. The seller obtains a revenue $\Pi^D(\mu_0) = t^D(\mu_0)$.

Proposition 2. *Under Learning Deterrence, the buyer's trading surplus $\mathbb{E}[v|\mu_0] - t^D(\mu_0)$ is non monotone and single-peaked in $\mu_0 \in [\underline{\mu}, \overline{\mu}]$. Moreover, $\mathbb{E}[v|\mu_0] - t_b^D(\mu_0) = 0$ at the two endpoints. The seller's revenue $\Pi^D(\mu_0)$ is increasing in μ_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.

A **Free Return** mechanism $\{t_b, (0,0)\}$, with $s \in [q^{-1}(\mu_0), Q^{-1}(\mu_0)]$, can implement the boundaries that the shaded area intersects with $q(s)$ and thereby encourage the buyer to learn. 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 implement different quitting beliefs and thereby induce different amounts of information acquisition which eventually affects the ex-ante expected probability of a successful sale. A common feature of Free Return is that the buyer stops learning when the continuation value from learning is 0. We call it as *full learning* since it is the largest amount of information acquisition that the seller can implement when the price is fixed.

Suppose the seller allows a free return. The constrained optimization problem (F)

[10]The value of information refers to the difference between the value function and the maximum payoff from purchasing and walking away.

below determines the optimal price.

$$\Pi^{\mathscr{F}}(\mu_0) := \max_{s\in[q^{-1}(\mu_0),Q^{-1}(\mu_0)]} \quad \frac{\mu_0 - q(s)}{1-q(s)}(v_h - s) \tag{F}$$

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 $\Pi^F(\mu_0)$. We can verify that $\Pi^{\mathscr{F}}(\mu_0) = \Pi^F(\mu_0)$ if $\Pi^{\mathscr{F}}(\mu_0) \geq \Pi^D(\mu_0)$.

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 $\mu_0 \in [\underline{\mu}, \overline{\mu}]$,

$$\Pi^*(\mu_0) \geq \max\left\{\Pi^D(\mu_0), \Pi^F(\mu_0)\right\}.$$

5 Stochastic Return

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 call *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 probability at return in order to compensate the buyer's continuation value from learning. Therefore, Stochastic Return guarantees a strictly positive minimum revenue as the buyer is willing to accept a partial refund in exchange for a positive odds of keeping the item. However, in this section, we show that Stochastic Return is sub-optimal.

We first derive the implementable mechanisms that can induce the interior points in the shaded area.

Lemma 2. *For fixed $s \in [\underline{s}, \bar{s}]$, the return policy $(x_r(\mu,s), t_r(\mu,s))$ implements stopping belief $\mu \in [q(s), Q(s)]$, where*

$$x_r(\mu,s) = \frac{V_1(\mu,s)}{v_h - v_l}, \tag{7}$$

$$t_r(\mu,s) = \mathbb{E}[v|\mu]x_r(\mu,s) - V(\mu,s). \tag{8}$$

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.*

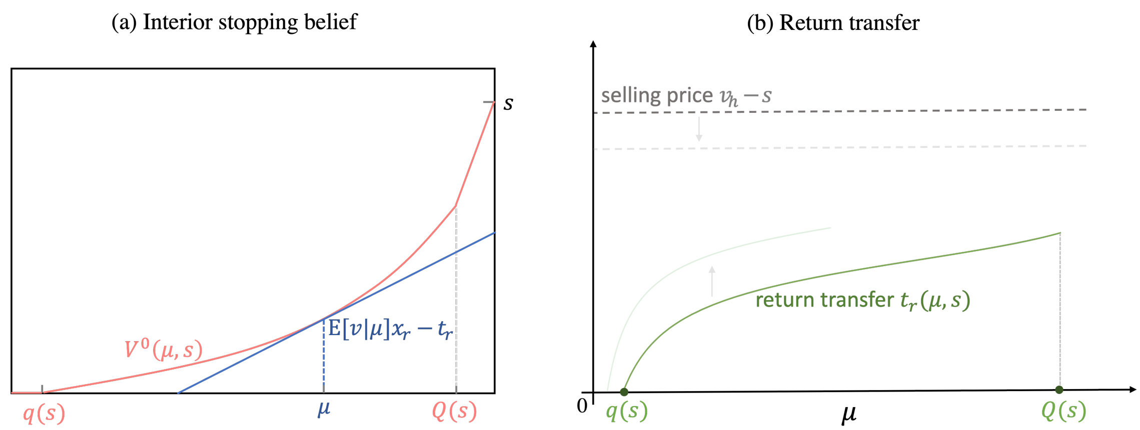


Figure 2: Stochastic Return

The allocation probability at return x_r is proportional to the slope of the buyer's value function. And t_r equals the allocation surplus upon return minus the continuation value from learning that the buyer holds when he decides to stop learning. A refund mechanism with price $v_h - s$ and return policy $(x_r(\mu,s), t_r(\mu,s))$ implements the stopping belief μ as the buyer stops learning at the belief he reports to the seller. Consider Figure 2(a). To induce the buyer to stop learning earlier at a belief higher than $q(s)$, the return policy must generate a payoff weakly higher than the buyer's continuation value from learning. Meanwhile, to make stopping learning at such belief incentive compatible, the return payoff is tangent to the continuation value from learning. In other words, equations (7) and (8) come from the familiar smooth pasting and value matching conditions.

Figure 2(b) plots the return transfer for a fixed s. Interestingly, $t_r(\mu,s)$ increases with μ, meaning that the seller actually obtains a larger return transfer if she tries to force the buyer to stop learning earlier when there is still a higher continuation value from learning. This is because, the seller also use x_r for the buyer compensation, and it turns out the seller allows the buyer to keep the item with a larger probability when she tries to implement a larger stopping belief. With similar reasoning, $t_r(\mu,s)$ 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. From Figure 2, we observe that,

$$\begin{aligned} \lim_{\mu\to q(s)} x_r(\mu,s) = 0 \qquad &\text{and} \qquad \lim_{\mu\to q(s)} t_r(\mu,s) = 0; \\ \lim_{\mu\to Q(s)} x_r(\mu,s) \ll 1 \qquad &\text{and} \qquad \lim_{\mu\to Q(s)} t_r(\mu,s) \ll v_h - s. \end{aligned} \tag{9}$$

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, because the return transfer $t_r(Q(s),s)$ is smaller than the price $v_h - s$.

5.1 Bang-Bang solution

We rewrite seller's optimization for encouraging learning as a maximization over the inducible set of (s,μ).

$$\max_{s\in[q^{-1}(\mu_0),Q^{-1}(\mu_0)]} \left\{ \begin{aligned} \max_{\mu} \quad & \Pi(\mu,s) := \frac{\mu_0-\mu}{1-\mu}(v_h - s) + \frac{1-\mu_0}{1-\mu} t_r(\mu,s) \\ \text{s.t.} \quad & q(s) \le \mu \le Q(s) \\ & \mu \le \mu_0 \end{aligned} \right\} \tag{L}$$

Conditional on acquiring information is optimal for type-μ_0 buyer, the inner maximization derives the optimal stopping belief that maximizes the seller's expected revenue for every fixed s. The outer maximization derives the optimal s on the solution path of the inner maximization.

Theorem 1. *Under the optimal refund mechanism, $x_r \in \{0,1\}$.*

That is, $\Pi^*(\mu_0) = \max\left\{\Pi^D(\mu_0), \Pi^F(\mu_0)\right\}$. The idea of the proof is to show that the solution locates on the boundary of the inducible learning outcomes, i.e., the boundary of shaded area. Consider Figure 3 for an illustration. It gives an example that $\mu_0 = 0.5$. The red curve (both dashed part and solid part) plots the solution to the inner maximization, i.e., the optimal stopping belief as a function of s. There is a cutoff $q^{-1}(\underline{\mu}^*)$ such that if $s > q^{-1}(\underline{\mu}^*)$, then the optimal stopping belief is an interior solution. Otherwise, if $s \le q^{-1}(\underline{\mu}^*)$, the optimal stopping belief is bounded by the quitting belief. Notice that the lower boundary of the solid red curve refers to a Free Return mechanism, while the upper boundary of the solid red curve is strictly dominated by Learning Deterrence.[11] Therefore, if the optimal refund mechanism induces interior outcome, then the outcome induced must locate on the interior path of the solid red curve. However, we show that the seller's revenue is either *quasi-convex* along the solid red curve or it adopts an interior local maximum which is dominated by Learning Deterrence. Thus, we establish the sub-optimality of Stochastic Return. 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.

[11] Notice that we can use both Stochastic Return and Learning Deterrence to induce the orange dot. However, the revenue from Learning Deterrence is strictly higher (implied by the second half of (9)).

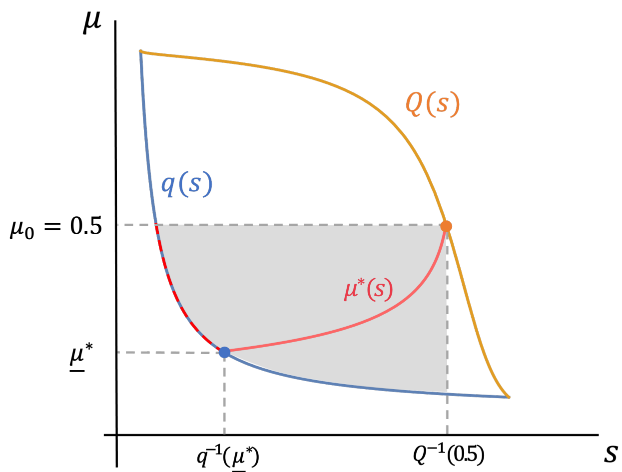


Figure 3: Stochastic Return and interior solutions

One more lesson from this exercise is that when the seller optimally allows a free return, she escalates price discrimination. Recall from Figure 3, Free Return can be optimal only when $s \leq q^{-1}(\underline{\mu}^*) \ll Q^{-1}(0.5)$. Therefore, the optimal price for Free Return $v_h - s^F(\mu_0)$ is much higher than the price for Learning Deterrence $v_h - Q^{-1}(0.5)$. We provide a detailed discussion of the proof in the remaining part of this section.

5.1.1 Inner Maximization

For fixed s, consider the inner maximization of (L) subject to $\mu \in [q(s), Q(s)]$,

$$\max_{\mu \in [q(s), Q(s)]} \quad \Pi(\mu, s). \tag{10}$$

Let $\mu^*(s) := \{\mu \in [q(s), Q(s)] : \Pi_1(\mu, s) = 0\}$ be the interior maximizer of (10). Rearranging $\Pi_1(\mu, s) = 0$, we obtain[12]

$$\underbrace{\Pr(\text{return}) \frac{\partial t_r(\mu, s)}{\partial \mu}}_{\text{larger return transfer}} = \underbrace{[v_h - s - t_r(\mu, s)] \frac{d\Pr(\text{return})}{d\mu}}_{\text{more frequent return}}. \tag{11}$$

Thus, the seller faces a trade-off between increasing the buyer's stopping belief so that to gain a larger return transfer and decreasing the buyer's stopping belief so that to decrease the expected return rate. The first order condition derives a unique path of $\mu^*(s)$. We can further obtain the monotonicity of $\mu^*(s)$. To see this, recall that Lemma 2 establishes

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

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 expected return rate, and her incentive to gain a larger return transfer is relatively stronger. Thus, she optimally adapts to implement 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 implement 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 expected return rate, and the seller implements the minimal stopping belief, rendering the lower boundary $q(s)$ the optimal stopping belief. We can further verify that $\mu^*(s)$ is independent of the prior belief μ_0.

Lemma 3. *Let $\underline{\mu}^*$ be the solution of $\Pi_1(\mu, q^{-1}(\mu)) = 0$. Then $\underline{\mu}^* < 0.5$.*
(1) If $s \le q^{-1}(\underline{\mu}^*)$, the optimal stopping belief is $q(s)$;
(2) If $s \in (q^{-1}(\underline{\mu}^*), Q^{-1}(0.5))$, the optimal stopping belief is $\mu^*(s)$;
(3) If $s \ge Q^{-1}(0.5)$, the optimal stopping belief is μ_0.

Lemma 3 summarizes the optimal stopping belief as s varies.[13] The second term of this lemma indicates that partial learning can be optimal if the value of s is intermediate. Inversely, given that $\mu^*(s)$ is strictly increasing in s, Stochastic Return can be optimal only when the optimal stopping belief $\mu \in (\underline{\mu}^*, 0.5)$, shown as the solid red curve in Figure 3.

5.1.2 Outer Maximization

The seller's profit along the solid red curve $\mu^*(s)$ equals:

$$\Pi(\mu, s^*(\mu)) = t_r(\mu, s^*(\mu)) + \frac{\mu_0 - \mu}{1-\mu}\Big[v_h - s^*(\mu) - t_r(\mu, s^*(\mu))\Big]. \tag{12}$$

where $s^*(\mu)$ represents the inverse of $\mu^*(\cdot)$ for $\mu \in [\underline{\mu}^*, 0.5]$. Note that the first term is the seller's minimum revenue, and the second term refers to the extra revenue she can obtain conditional on that the buyer discovers good news.

We show that seller's profit is either quasi-convex along the path of $\mu^*(s)$, or there is a local maximizer on $\mu^*(s)$ which is dominated by Learning Deterrence. Interestingly, under the case that $v_l = 0$, the former case is always true.[14] Intuitively, suppose that

[13]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 $\mu = 0.5$.

[14]For $v_l > 0$, if the learning cost is not very low, then seller's profit is still quasi-convex along $\mu^*(s)$.

the seller tends to lower the buyer's stopping belief to induce more extended learning process, 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 implement a lower stopping belief, the minimum revenue $t_r(\mu, s^*(\mu))$ also decreases,[15] causing the extra revenue gain from a successful sale 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 of $\mu^*(s)$. Conversely, suppose the seller tends to implement 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 3). Therefore, Stochastic Return is sub-optimal.

6 Optimal Refund Mechanism

Recall that Theorem 1 implies $\Pi^*(\mu_0) = \max\left\{\Pi^D(\mu_0), \Pi^F(\mu_0)\right\}$. Denote the set of prior beliefs that the seller weakly prefers Free Return as $F := \left\{\mu_0 \in [\underline{\mu}, \overline{\mu}] : \Pi^F(\mu_0) \geq \Pi^D(\mu_0)\right\}$.

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 a price $\mathbb{E}[v|\mu_0]$ if $\mu_0 \notin [\underline{\mu}, \overline{\mu}]$;*
2. *Learning Deterrence (no return) with a price $t^D(\mu_0)$ if $\mu_0 \in [\underline{\mu}, \overline{\mu}]$ and $\mu_0 \notin F$;*
3. *Free Return with a ;price $v_h - s^F(\mu_0)$ if $\mu_0 \in F$.*

We can interpret $\frac{k}{\lambda}$ as the effective learning cost. Figure 4 depicts the expected revenue/price 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, there exist two cutoff beliefs such that the seller optimally chooses Free Return when the prior belief lies in between. Otherwise, the seller optimally chooses Learning Deterrence.

To interpret this result, note that 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

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

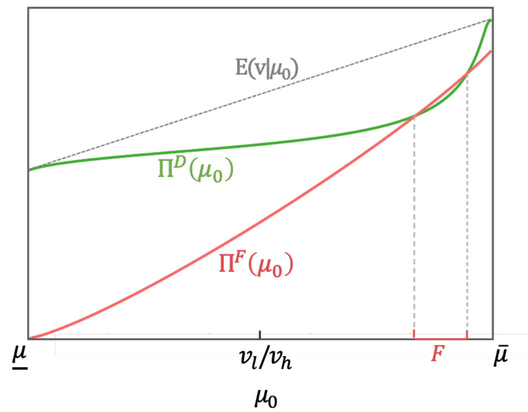
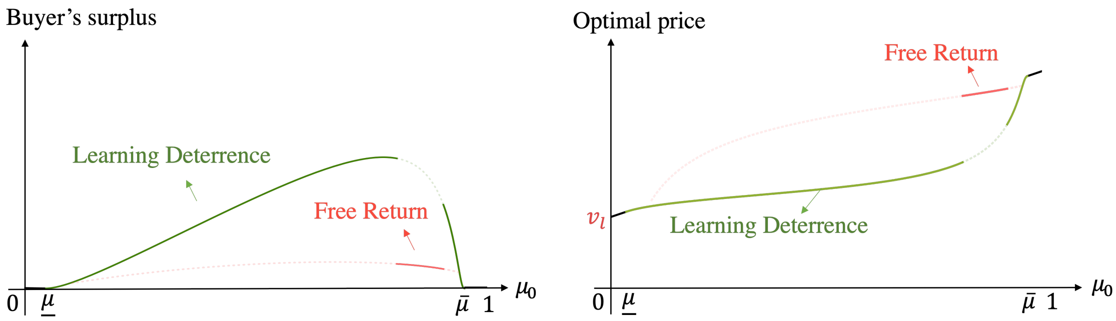


Figure 4: Learning Deterrence and Free Return revenue for small learning cost

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 significantly reduce the price to prevent learning, rendering Learning Deterrence less profitable. However, if the seller switches to Free Return to encourage learning, she can avoid compensating the buyer's opportunity information rent. That is, instead of significantly decreasing 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 a successful sale.

Essentially, the buyer's prior belief measures (1) how much the buyer values information ex-ante; and (2) how optimistic the buyer initially is. Thus, 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.

The left panel of Figure 5 plots the buyer's expected trading surplus against the prior belief under the optimal mechanism. There is a severe decline when the seller optimally allows Free Return, which is driven by price discrimination under Free Return (see the right panel for the price of the optimal refund mechanism).



(a) buyer's expected surplus

(b) Optimal selling price

Figure 5: Buyer's surplus and selling price

When the effective learning cost is large, $\frac{k}{\lambda} > \gamma^*$, the buyer obtains little value from learning. Therefore, the price reduction the seller has to offer to deter learning is insignificant 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 prior beliefs F that supports Free Return as the optimal mechanism shrinks; and eventually becomes an empty set when $\frac{k}{\lambda} > \gamma^*$. See the bottom row in Figure 6.

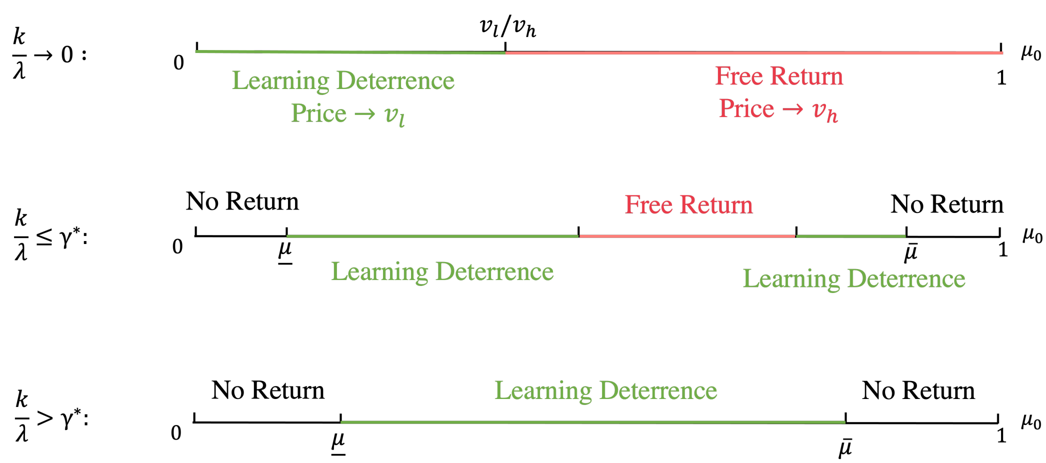


Figure 6: Optimal refund mechanism

Conversely, when learning becomes less costly, the set of prior beliefs F that supports Free Return expands. As the effective learning cost converges to zero, the buyer can learn almost perfect information. Therefore, to deter learning, 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 (see the first row in Figure 6).

Proposition 3. $\Pi^F(\mu_0)$ *is decreasing in the effective learning cost, while* $\Pi^D(\mu_0)$ *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. When the effective learning cost converges to zero, the solution converges to standard screening solution with perfectly informed buyer.*

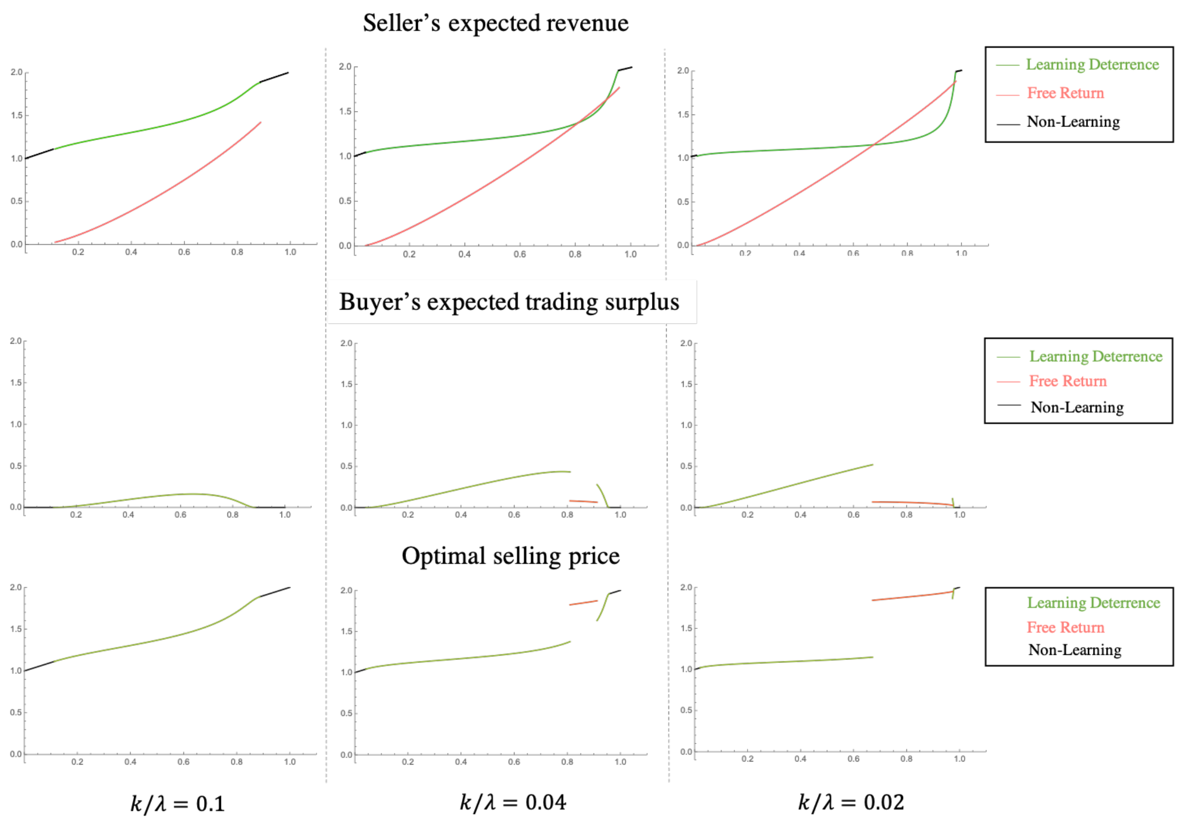


Figure 7: Comparative statics

Figure 7 depicts the seller's revenue, the optimal selling price and the buyer's expected trading surplus while the effective learning cost $\frac{k}{\lambda}$ goes down. Consider the second row from the left to the right. The information cost has a non-monotone effect in terms of the buyer's expected trading surplus. This is because the buyer can only benefit from lower information cost if the seller deters learning. However, if the information cost becomes sufficiently low, the seller switches to Free Return and escalates price discrimination, which eventually hurts the buyer.

7 More Efficient Post-purchase Learning

If the learning process is more efficient after a transaction, i.e., $\lambda_P > \lambda_B$, then 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 4. *If the optimal mechanism allows a return, the cancellation fee t_u is the solution to the following equation,*

$$V(\mu_0, s + t_u; \lambda_P) - t_u = V^0(\mu_0, s; \lambda_B).^{16} \tag{13}$$

The left hand side is the buyer's continuation value for purchasing under the optimal refund mechanism, while the right hand side is the continuation value if a return is not allowed. Proposition 4 implies that, for any optimal refund mechanism, the buyer obtains the same continuation value as if the mechanism prohibited return. If $\lambda_P = \lambda_B$, then $t_u = 0$. Proposition 4 generalizes Lemma 1.

If the optimal mechanism deters buyer learning, then it takes the same form as Learning Deterrence, no return with price $t_b = t^D(\mu_0; \lambda_B)$, 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

[16] $V(\mu_0, s + t_u; \lambda_P)$ is the solution to (ODE) where we substitute λ by λ_P and impose a boundary point $(q(s + t_u; \lambda_P), 0)$.

Lemma 2. In particular,

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

where the allocation probability at return is proportional to the slope of buyer's value function for post-purchase learning, and the return transfer equals the allocation surplus at return minus the continuation value from learning. Thus, to encourage learning, the seller's optimization problem is the following.

$$\max_{s\in[q^{-1}(\mu_0;\lambda_P),Q^{-1}(\mu_0;\lambda_B)]} \left\{ \max_{\mu} \frac{\mu_0-\mu}{1-\mu}(v_h - s) + \frac{1-\mu_0}{1-\mu}\left(t_r(\mu,s)+t_u\right)\right\} \tag{14}$$
$$\text{s.t.} \quad q(s+t_u;\lambda_P) \le \mu \le Q(s;\lambda_B)$$
$$\mu \le \mu_0$$

If λ_P is close to λ_B, deterministic mechanism is still optimal, i.e., $x_r \in \{0,1\}$, since the cancellation fee is not very large. However, $x_r = 0$ implies that there is zero allocation surplus if the seller matches with a low valuation buyer, which is inefficient because 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 5. *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:*

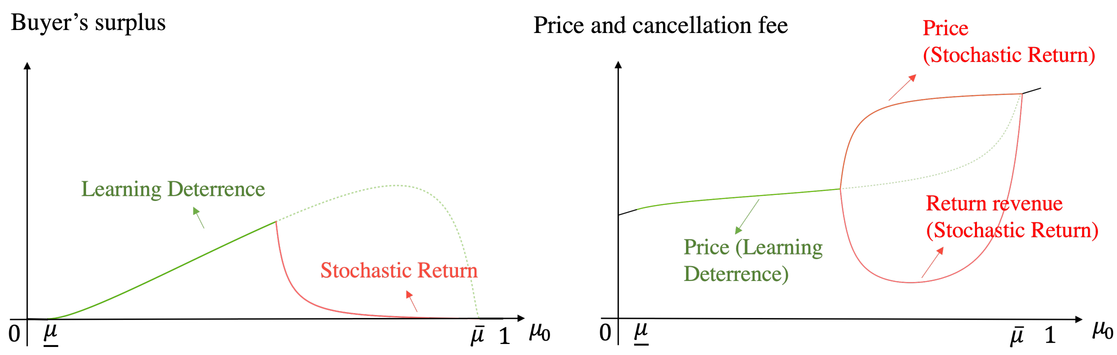
$$t_b = t^D(\mu_0;\lambda_B),\ t_u = 0,\ \text{and}\ (x_r,t_r) = (1,t_b);$$

2. *Stochastic Return:*

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

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 x_r and sets $t_r = x_r v_l$ to extract the return allocation surplus if matching with a low valuation buyer. Furthermore, she charges a cancellation fee to extract buyer's extra information rent from post-purchase learning.

In Figure 8, 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.



(a) buyer's expected surplus

(b) Price and cancellation fee

Figure 8: Buyer's surplus and optimal selling price if $\lambda_P \to \infty$

As the seller uses Stochastic Return to mitigate the efficiency loss when matching with a low valuation buyer, such incentive is gone if $v_l = 0$ and deterministic mechanism is still optimal.

Corollary 2. *If $\lambda_P \to \infty$ and $v_l = 0$, then one of the mechanisms below is optimal:*

1. *Learning Deterrence: no return with a price $t^D(\mu_0; \lambda_B)$;*
2. *Return with a cancellation fee: $t_b = v_h - \frac{k}{\lambda_B \mu_0}$, $x_r = 0$, $t_r = 0$ and $t_u = \frac{k}{\lambda_B(1-\mu_0)}$.*

8 No News is Good 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 remain the same.

The key difference between positive learning and negative learning is that, under positive learning, the 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 this case. Moreover, if we denote $\eta \equiv x_r v_l - t_r$ as the buyer's surplus while requesting a return upon observing bad news, then under every optimal mechanism,

$$\eta \equiv x_r v_l - t_r = 0 \Longrightarrow x_r = \frac{t_r}{v_l}.$$

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, $V^N(\mu, t_b)$. 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 $g(t_b)$ is determined by the indifference between returning the product and continuing to learn,

$$g(t_b) = \{\mu : \mathbb{E}[v|\mu]\frac{t_r}{v_l} - t_r = V^N(\mu, t_b)\}.$$

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,

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

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

Denote $G^{-1}(\mu) := v_l + \frac{k}{\rho(1-\mu)}$ as the inverse function of $G(t_b)$. Thus, the seller can implement a stopping belief 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 (upper) stopping belief that is implementable given the prior belief μ_0. We can then formulate the seller's optimization problem (N) as following.

$$\begin{aligned} \max_{\mu} \quad & \Pi^N(\mu) := \frac{\mu_0}{\mu}G^{-1}(\mu) + \frac{\mu - \mu_0}{\mu}t_r \\ \text{s.t} \quad & V^N(\mu_0, G^{-1}(\mu)) = \mathbb{E}[v|\mu_0]\frac{t_r}{v_l} - t_r, \\ & \mu_0 \leq \mu \leq \bar{G}(\mu_0). \end{aligned} \tag{N}$$

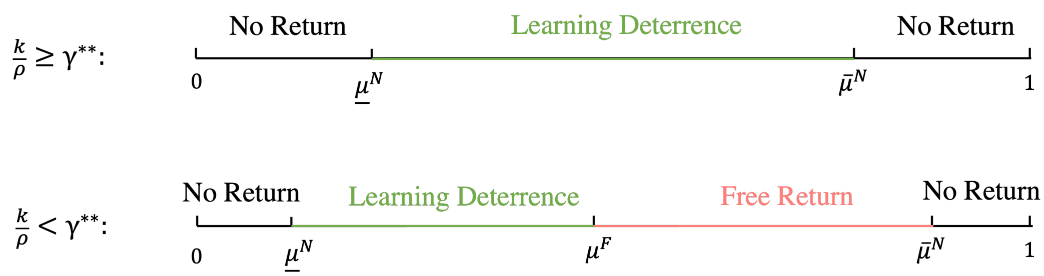


Figure 9: Optimal refund mechanism with bad news

Similarly, the seller's expected revenue is a weighted average between the selling price and the return transfer. Notice that $g(t_b)$ does not affect the ex-ante probability of a successful sale or a return. Therefore, for any fixed price, the seller can always obtain a higher return transfer by increasing allocation probability at return, which then induces a higher $g(t_b)$. It implies that the lower stopping belief equals the prior belief under the optimal refund mechanism, which is the first constraint. The second constraint characterizes the implementable consumption belief. Note that unlike the program under positive learning, the objective function is continuous at $\mu = \mu_0$ if we switch from encouraging learning to deterring learning. Therefore, we no longer require two programs to study encouraging learning and deterring learning respectively. 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,

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

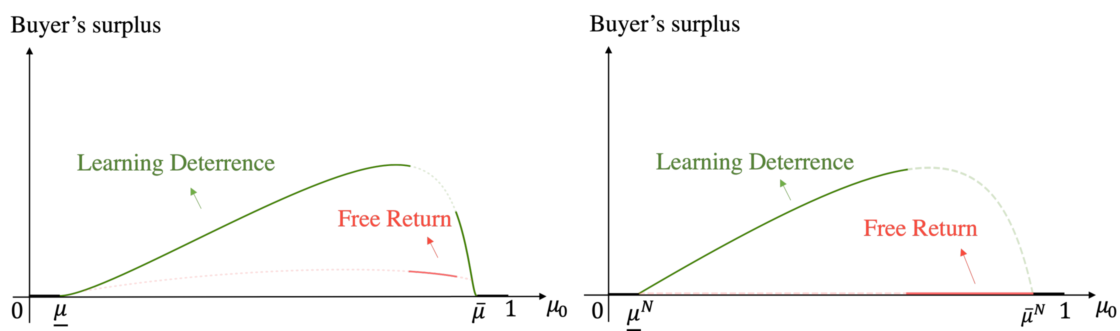
Proposition 6. *Assume $4k < (v_h - v_l)\rho$. There exists a γ^{**} such that if $k/\rho < \gamma^{**}$, then $\mu^F < \overline{\mu}^N$ and the optimal mechanism takes following form:*

1. *No Return with a price $\mathbb{E}[v|\mu_0]$ if $\mu_0 \notin [\underline{\mu}^N, \overline{\mu}^N]$;*
2. *Learning Deterrence (no return) with a price $G^{-1}(\mu_0)$ if $\mu_0 \in [\underline{\mu}^N, \mu^F]$;*
3. *Free Return with a price $G^{-1}(\bar{G}(\mu_0))$ 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.

The optimal mechanism under negative learning (described in Proposition 6) takes a similar form as under positive learning (described in Theorem 2). However, the right cutoff prior belief that the seller optimally chooses Free Return equals $\overline{\mu}^N$, shown as the second case in Figure 9. Intuitively, if $\mu_0 = \overline{\mu}^N$, the largest 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. 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 10 gives a comparison of buyer's ex-ante surplus under positive learning and negative learning respectively. In the right panel, under negative learning, the buyer can only obtain positive participation value if the mechanism deters learning.



(a) Positive learning (b) Negative learning

Figure 10: The buyer's ex-ante surplus

9 Discussion

There is an alternative interpretation of the refund mechanism we study in this paper. Consider a start-up company offering an option contract to one acquirer. The option contract contains a baseline allocation and an option. The baseline allocation specifies the price t_r for a share x_r of the company. On top of it, the acquirer has an option to purchase the remaining share $(1 - x_r)$ of the company at a strike price $(t_b - t_r)$. The timeline is the following. At the beginning, the start-up offers this mechanism to the acquirer. The acquirer can then evaluate this contract before contracting with the start-up. If the acquirer decides to contract with the start-up, then based on the information arrives afterwards, he can decide whether to purchase the remaining share. Notice that if the baseline allocation specifies the price for the entire share, i.e., $x_r = 1$ and $t_b - t_r = 0$, then the contract does not involve any real option.

With this interpretation, in the benchmark model under "no news is bad news", the buyer purchases the remaining share only if he obtains a conclusive good news before he stops learning, e.g., some ongoing innovation of the start-up is proved to be success. Anticipating this, the start-up designs the baseline allocation to control the buyer's learning decision, as the buyer decides to stop learning comparing the expected payoff from the baseline allocation and the continuation value from learning which also depends on the strike price. Given Theorem 1, the optimal baseline allocation either ensures the buyer an option to freely opt out, or specifies a price for the entire share. Moreover, in the former case, the start-up designs a sufficiently high strike price such that the buyer is willing to learn for a relatively long time, while in the latter case, the price is sufficiently low so that the buyer is willing to purchase the entire company immediately. The optimal contract has a similar structure under "no news is good news" framework.

Regulations on buyer's right for free opt-out are designed presumably for the sake of buyer's surplus. However, it actually diminishes the buyer's expected surplus as it triggers price discrimination. Specifically, when learning is inevitable on the buyer side, the seller then has a strong incentive to give up those buyers receiving bad news but enhances price discrimination to those buyers receiving good news. One possible solution is to impose a price cap, which then limits the room for price discrimination. Conditional on Lemma 3, a price cap might eventually make the seller optimally choose the deterring learning price (much lower than the price cap) and create efficient allocation.

Appendix

Proof of Lemma 1. Suppose (IR) holds with strict inequality. That is, the seller offers a benevolent return policy $\{x_r(\cdot), t_r(\cdot)\}$ such that if the buyer reports a belief μ and optimally stops post-purchase learning at belief μ and requests a return, he gets a payoff $V_P(\mu; t_b, x_r(\mu), t_r(\mu)) > V^0(\mu; t_b, 1, t_b)$. Given (IM-L2)—the optimality to stop learning at μ, we can then calculate the return transfer $t_r(\mu)$ which equals the allocation surplus at return minus the buyer's continuation value:[17]

$$\begin{aligned} t_r(\mu) &= \mathbb{E}[v|\mu]x_r(\mu) - V_P(\mu; t_b, x_r(\mu), t_r(\mu)) \\ &= \mathbb{E}[v|\mu]\frac{V_P'(\mu; t_b, x_r(\mu), t_r(\mu))}{v_h - v_l} - V_P(\mu; t_b, x_r(\mu), t_r(\mu)). \end{aligned} \tag{15}$$

Note that $V_P(\mu; t_b, x_r(\mu), t_r(\mu))$ must satisfy the differential equation,

$$(1-\mu)\mu\lambda V_P'(\mu; t_b, x_r(\mu), t_r(\mu)) + \mu\lambda V_P(\mu; t_b, x_r(\mu), t_r(\mu)) = \mu\lambda(v_h - t_b) - k,$$

derived from the Bellman equation for post-purchase learning (3). Slope V_P' and magnitude V_P of the buyer's continuation value are the substitutes that the seller can adjust to implement the same stopping belief. For the purpose of maximizing profit, the seller can reduce $V_P(\mu; t_b, x_r(\mu), t_r(\mu))$ and raise $V_P'(\mu; t_b, x_r(\mu), t_r(\mu))$ without violating the above differential equation, which increases the return transfer and in the meantime preserve the same buyer's optimal stopping rule. Moreover, by implementing the same stopping beliefs, the ex-ante probabilities of return and successful sale are the same. This implies a profitable deviation. □

Proof of Proposition 1. 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.*

[17] This is implied by the optimality (known by smooth-pasting and value matching conditions) to stop learning at μ.

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

$$V(\mu, D(\mu)) = \mathbb{E}[v|\mu] - (v_h - D(\mu)). \tag{16}$$

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

$$\frac{dD(\mu)}{d\mu} = \frac{k(k-\lambda D(\mu))}{\lambda^2(1-\mu)^2\mu D(\mu)} = \frac{k[\tilde{\mu}-1]}{\lambda(1-\mu)^2\mu} < 0. \tag{17}$$

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 $(\underline{\mu},\underline{\mu})$,[18] and its solution is,[19]

$$-\frac{1}{\tilde{\mu}} - \log[1-\tilde{\mu}] + \log[\tilde{\mu}] = \frac{1}{1-\mu} - \log[1-\mu] + \log[\mu] - \frac{\lambda(v_h - v_l)}{k}. \tag{18}$$

Denote the LHS as $f(\tilde{\mu})$ and the RHS as $g(\mu)$. The domain of both functions is $[\underline{\mu},\overline{\mu}]$ and $f(\cdot) = g(\cdot)$ at the two end points. Note that $f'(\cdot) > g'(\cdot)$ when the both arguments are smaller then 0.5 and $f'(\cdot) < g'(\cdot)$ when both arguments are larger then 0.5.[20] Therefore $f(\cdot)$ and $g(\cdot)$ cross only at the two boundary points and therefore $\tilde{\mu}(\mu) < \mu$ for all $\mu \in (\underline{\mu},\overline{\mu})$. For $\tilde{\mu}(\mu)$ to be symmetric about $1-\mu$, note that the reflection point of $(\mu,\tilde{\mu})$ over line $1-\mu$ is $(1-\tilde{\mu}, 1-\mu)$. It is easy to verify that, if equation (18) holds at a point $(\mu,\tilde{\mu})$, then equation (18) 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 $\tilde{\mu}(\mu) = 1-\mu$.[21] 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. □

When $s \notin [\underline{s},\bar{s}]$, as constraint (6) 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

[18] To verify $(\underline{\mu},\underline{\mu})$ is an initial point. Recall $\underline{\mu} = q(\bar{s})$ and the binding (6) 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 (16), $D(\underline{\mu}) = \bar{s}$. Thus, $\tilde{\mu}(\underline{\mu}) = q(D(\underline{\mu})) = \underline{\mu}$.

[19] 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.

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

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

large s, and $Q(s) = q(s)$ at $\underline{s}$ and $\bar{s}$.[22] 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. □

Proof of Proposition 2. First, we prove the first term. Proposition 1 and the constraint (6) imply $V(\mu, D(\mu)) = 0$ at $\underline{\mu}$ and $\bar{\mu}$. Rearranging equation (16) gives:

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

Taking derivative w.r.t μ and plugging in equation (17) 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 $A = \frac{k}{\lambda(v_h - v_l)} = (1-\underline{\mu})\underline{\mu} \in (0, \frac{1}{4})$.[23] 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 that $\frac{dV(\mu, D(\overline{\mu}))}{d\mu} = 0$ 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 (18).

$$-A\frac{(1-\tilde{\mu})}{(1-\mu)^2\mu} + 1 = 0 \tag{19}$$

$$-\frac{1}{\tilde{\mu}} + \log\left[\frac{\tilde{\mu}}{1-\tilde{\mu}}\right] = \frac{1}{1-\mu} + \log\left[\frac{\mu}{1-\mu}\right] - \frac{1}{A} \tag{20}$$

Substituting equation (19) into (20), 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

[22] Recall that $D(\underline{\mu}) = \bar{s}$ and $D(\bar{\mu}) = \underline{s}$ by (6).

[23] From the binding (6), 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, $A \in (0, \frac{1}{4})$.

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 $(1-\mu)^2\mu - A = 0$.[24] We can show that $z(\mu)$ crosses $y(\mu)$ twice in the support $[\underline{\mu}, \overline{\mu}]$, first from above and then from below.[25]

Next, the monotonicity of $\Pi^D(\mu_0) = t^D(\mu_0) = v_h - D(\mu_0)$ can be directly obtained from (17). 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 $t^D(\overline{\mu}) = \mathbb{E}(v|\overline{\mu})$. □

Proof of Lemma 2. Given Lemma 1, to induce the buyer to stop learning at a belief μ different from $q(s)$, the buyer's expected payoff from requesting return $\mathbb{E}(v|\cdot)x_r - t_r$ should smoothly pass $V^0(\cdot, s)$ at μ. Besides, the induced stopping belief μ must belong to the set $[q(s), Q(s)]$, in which $V^0(\mu, s) = V(\mu, s)$. 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 in (7) and (8). Specifically,

$$t_r(\mu, s) = -\frac{kv_l - \lambda\mu v_l s - k\mu v_h\left[\log\left(\frac{\mu}{1-\mu}\right) - \log\left(\frac{k}{\lambda s - k}\right)\right]}{\lambda\mu(v_h - v_l)}. \tag{21}$$

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

[25] (1) Suppose $A < 4/27$, then $z(\mu) > y(\mu)$ for $\underline{\mu} \leq 1/3$, $z(r_2) = 0 < y(r_2)$ 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 $y(\overline{\mu}) < z(\overline{\mu})$.

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. □

Proof of Lemma 3. First, we discuss the first-order condition. Explicitly,

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

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 $\mu^*_-(s) \le \mu^*_+(s)$. 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, $\mu^*_-(s) = \mu^*_+(s) > 0.5$.[26] Thus, if there exists a $\mu^*_+(s)$, it is larger than 0.5.

Let $s^*(\mu) = \{s : \Pi_1(\mu,s) = 0\}$.[27] 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 $\mu \le 0.5$, $s^*(\mu)$ 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,

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

[26] 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.

[27] 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.

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.

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 $t_r(\mu_+^*(s),s) > \bar{t}_r(\mu_+^*(s))$,[28] 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 find the solution to the system of equations below,

$$\begin{cases} \Pi_1(\mu,s) = 0, \\ V(\mu,Q^{-1}(\mu)) = \mathbb{E}[v|\mu] - (v_h - Q^{-1}(\mu)), \\ s = Q^{-1}(\mu), \end{cases}$$

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

[28] Note that equation (21), the exact expression of $t_r(\mu,s)$, is valid for all $\mu \in [0,1]$.

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 $\underline{\mu}^* = \frac{k}{\lambda v_h} + (\frac{k}{\lambda v_h}(\frac{k}{\lambda v_h} + \frac{v_l}{v_h - v_l}))^{\frac{1}{2}}$.[29] Therefore, we pin down the set $[\underline{\mu}^*, 0.5]$ on which $t_r(\mu, s^*(\mu)) \in [0, \bar{t}_r(\mu)]$. □

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 $\mu \le 0.5$, 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 $s \ge Q^{-1}(0.5)$, then $Q(s) \le 0.5 \le \mu_-^*(s)$, 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 implement a return belief $Q(s)$. If $s \le q^{-1}(\underline{\mu}^*)$, $q(s) \ge \underline{\mu}^*$ and then $\Pi_1(q(s), s) \le 0$.[30] Since $\Pi_1(\mu, s)$ is quasi-concave in $[q(s), Q(s)]$, thus if $\Pi_1(\mu, s) \le 0$ at $q(s)$, $\Pi_1(\mu, s) \le 0$ for all $[q(s), Q(s)]$. Still the inequality holds with equality only at $s = q^{-1}(\underline{\mu}^*)$. It is optimal to implement return belief $q(s)$. □

Proof of Theorem 1. Substituting the first-order condition (11) into the seller's expected revenue (12),

$$\Pi(\mu, s^*(\mu)) = t_r^*(\mu) + \frac{\partial t_r(\mu, s^*(\mu))}{\partial \mu}(\mu_0 - \mu). \tag{22}$$

Taking the derivative w.r.t μ gives

$$\frac{d\Pi(\mu, s^*(\mu))}{d\mu} = \frac{dt_r^*}{d\mu} - \frac{\partial t_r^*}{\partial \mu} + (\mu_0 - \mu)\frac{\partial^2 t_r^*}{\partial \mu^2} = \frac{\partial t_r^*}{\partial s}\frac{ds^*}{d\mu} + (\mu_0 - \mu)\frac{\partial^2 t_r^*}{\partial \mu^2} = \left[\frac{\frac{\partial t_r^*}{\partial s}\frac{ds^*}{d\mu}}{\frac{\partial^2 t_r^*}{\partial \mu^2}} + \mu_0 - \mu\right]\frac{\partial^2 t_r^*}{\partial \mu^2}$$

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

[29] 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.

[30] See footnote 29.

Note that $\frac{\partial t_r(\mu, s^*(\mu))}{\partial \mu}$ is independent of s and we can verify $\frac{\partial^2 t_r^*}{\partial \mu^2} < 0$.[31]

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. If $\frac{k}{\lambda} \geq \frac{v_h(\sqrt{v_h}-\sqrt{v_l})\sqrt{v_l}}{2v_h+(\sqrt{v_h}-\sqrt{v_l})\sqrt{v_l}}$, $\phi'(\underline{\mu}^*) \geq -1$, otherwise, $\phi'(\underline{\mu}^*) < -1$. Moreover, $\phi'(0.5) > -1$ for all $\frac{k}{\lambda} \in [0, \frac{v_h - v_l}{4}]$.*

The proof of this claim can be found subsequent to this theorem. Recall Lemma 3, $\mu^*(s)$ is an optimal solution only for the interval $s \in [q^{-1}(\underline{\mu}^*), Q^{-1}(0.5)]$. Consider the original problem (L) and reimpose the two constraints: $q^{-1}(\mu_0) \leq s \leq Q^{-1}(\mu_0)$ and $\mu \leq \mu_0$, then $\mu^*(s)$ is the 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

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

Figure 11 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 $\mu_0 \in [\underline{\mu}^*, 0.5]$, then the optimal stopping belief $\mu \leq \mu_0$ (see Figure 11 (a)); if $\mu_0 \in (0.5, Q(q^{-1}(\underline{\mu}^*))]$, then the optimal stopping belief $\mu \leq \mu^*(Q^{-1}(\mu_0))$ (see Figure 11 (b)), The lower boundary $\underline{\mu}^*$ can always be achieved when $\mu_0 \in [\underline{\mu}^*, Q(q^{-1}(\underline{\mu}^*))]$.

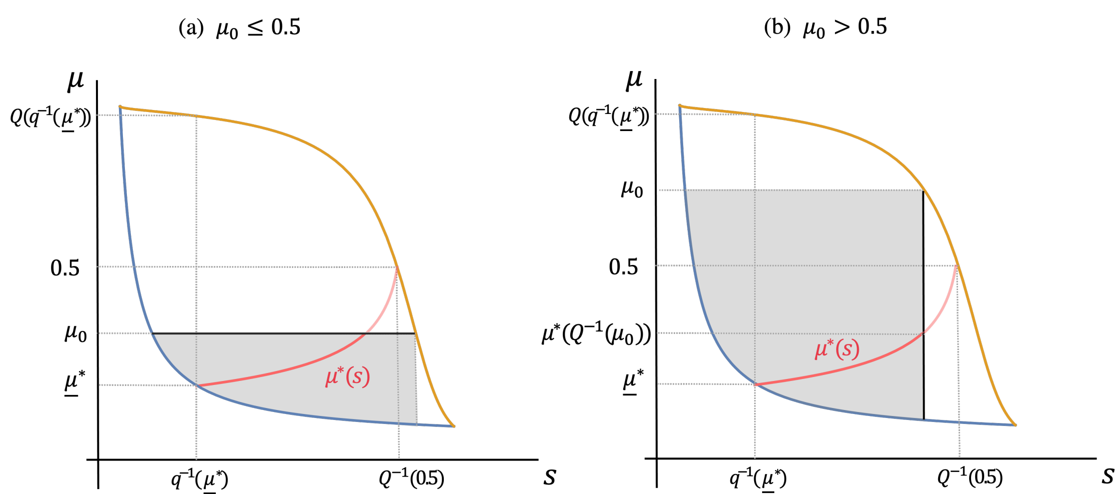


Figure 11: Feasible range of interior solutions

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

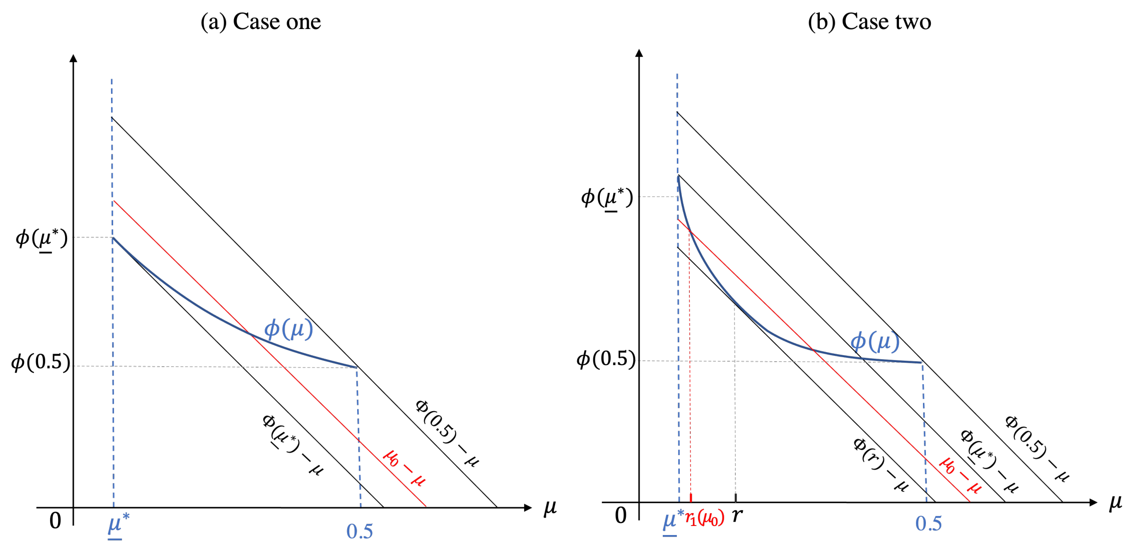


Figure 12: Illustration of the two cases

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. Notice that if $v_l = 0$, only case one is possible.

Case 1: $\phi'(\underline{\mu}^*) \geq -1$. 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 12(a).

Case 1(a), 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, the seller's revenue in this case is a weighted average between the deterring learning price $v_h - Q^{-1}(\mu_0)$ and the return transfer $t_r(\mu^*(Q^{-1}(\mu_0)), Q^{-1}(\mu_0))$, which is smaller than the deterring learning price. Case 1(b), if $\mu_0 \in [\Phi(\underline{\mu}^*), \Phi(0.5))$, $\Pi(\mu, s^*(\mu))$ is quasi-convex in μ. When $\mu_0 \leq 0.5$, 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. Case 1(c), if $\mu_0 \geq \Phi(0.5)$, $\Pi(\mu, s^*(\mu))$ 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 12(b)). If $\mu_0 \notin [\Phi(r), \Phi(\underline{\mu}^*)]$, then the expected revenue is quasi-convex and the argument in case one validates. If $\mu_0 \in [\Phi(r), \Phi(\underline{\mu}^*)]$, we want to show that

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

That is, the local maximum is not a global solution as it is worse than Learning deterrence. 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 (22) 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. □

Proof of Claim 3. 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 $\mu \in [\underline{\mu}^*, 0.5]$,[32] hence $w(\mu)$ is convex. 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.

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

Denote $q^*(\mu) := q(s^*(\mu))$. Simplifying $\phi'(\underline{\mu}^*) < -1$ implies,

$$(\underline{\mu}^*)^2(v_h - v_l) + 2\underline{\mu}^* v_l - v_l < 0.$$

Therefore, if $\underline{\mu}^* \in [0, \frac{\sqrt{v_l}}{\sqrt{v_h}+\sqrt{v_l}})$, then $\phi'(\underline{\mu}^*) < -1$. Otherwise if $\underline{\mu}^* \in [\frac{\sqrt{v_l}}{\sqrt{v_h}+\sqrt{v_l}}, 0.5)$, then $\phi'(\underline{\mu}^*) \geq -1$. Recall that $\underline{\mu}^* = \frac{k}{\lambda v_h} + (\frac{k}{\lambda v_h}(\frac{k}{\lambda v_h} + \frac{v_l}{v_h - v_l}))^{\frac{1}{2}}$ which is increasing in $\frac{k}{\lambda}$, where $\underline{\mu}^* \to 0$ if $\frac{k}{\lambda} \to 0$ and $\underline{\mu}^* \to 0.5$ if $\frac{k}{\lambda} \to \frac{v_h - v_l}{4}$. Thus, letting $\underline{\mu}^* = \frac{\sqrt{v_l}}{\sqrt{v_h}+\sqrt{v_l}}$ implies a unique cutoff of $\frac{k}{\lambda}$ such that if $\frac{k}{\lambda} < \frac{v_h(\sqrt{v_h}-\sqrt{v_l})\sqrt{v_l}}{2v_h+(\sqrt{v_h}-\sqrt{v_l})\sqrt{v_l}}$, $\phi'(\underline{\mu}^*) < -1$.

Moreover, 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 verify that $\phi'(0.5) > -1$ given that $q^*(0.5) < 0.5$. □

Proof of Theorem 2. Recall (F), the objective function is verified to be concave in s.[33] We solve the explicit solution for the unconstrained maximizer $s^F(\mu_0)$ and the unconstrained maximum $\Pi^F(\mu_0)$:

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

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

We first verify that, the constrained maximum $\Pi^{\mathscr{F}}(\mu_0) = \Pi^F(\mu_0)$ if $\Pi^{\mathscr{F}}(\mu_0) \geq \Pi^D(\mu_0) = t^D(\mu_0)$. It is equivalent to show that if $\Pi^{\mathscr{F}}(\mu_0) \geq t^D(\mu_0)$, then $q^{-1}(\mu_0) \leq s^F(\mu_0) \leq Q^{-1}(\mu_0)$. Obviously, $\Pi^{\mathscr{F}}(\mu_0) \geq t^D(\mu_0)$ implies $v_h - s^F(\mu_0) > t^D(\mu_0) = v_h - Q^{-1}(\mu_0)$, 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 $q^{-1}(\mu_0) \leq s^F(\mu_0)$, we want to show $q(s^F(\mu_0)) < \mu_0$, which can be simplified to $\sqrt{\frac{\mu_0}{1-\mu_0}} > \sqrt{\frac{k/(\lambda v_h)}{1-k/(\lambda v_h)}}$. It is true because $\frac{k}{\lambda v_h} < \underline{\mu} < \mu_0$.

Second, 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 $\Pi^F(\mu_0)$ intersects $t^D(\mu_0) = v_h - Q^{-1}(\mu)$ at most twice. That is,

$$Q^{-1}(\mu) = v_h - \Pi^F(\mu) \tag{23}$$

[33]The second order derivative w.r.t s is $\frac{2k\lambda(\mu_0-1)(k-\lambda v_h)}{(k-\lambda s)^3} < 0$.

has at most two roots when $\mu \in [\underline{\mu}, \overline{\mu}]$. Substituting (23) into equation (18), we obtain

$$g(\mu) - f(\theta(\mu)) = 0, \tag{24}$$

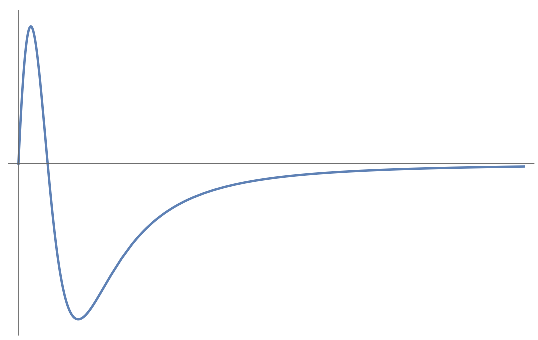
where $\theta(\mu) := \frac{\gamma}{(v_h - \Pi^F(\mu))}$ and $\gamma \equiv \frac{k}{\lambda}$. The first order derivative of the left hand side is

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

where $r = \sqrt{v_h/\gamma - 1} > \sqrt{3}$ given the assumption that $v_h > 4\gamma + v_l$. Let $x \equiv \sqrt{\frac{\mu}{1-\mu}} \in (0, \infty)$, 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 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 (24) has at most two roots.

Proposition 3 then implies the existence of γ^* and 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. □

Proof of Proposition 3. Recall that $\Pi(q(s), s) = \frac{\mu_0 - \gamma/s}{1-\gamma/s}(v_h - s)$. By the envelope theorem, we have:

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

To show that $t^D(\mu_0) = v_h - Q^{-1}(\mu_0)$ is increasing in γ, we want to show $Q^{-1}(\mu_0)$ is decreasing in γ. Recall that $\tilde{\mu}(\mu) := q(Q^{-1}(\mu))$. Taking the derivative w.r.t γ for both sides of $\mathbb{E}[v|\mu_0] - (v_h - Q^{-1}(\mu_0)) = V(\mu_0, Q^{-1}(\mu_0))$, we obtain:

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

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 $\mathbb{E}[v|\mu] - (v_h - q^{-1}(\mu)) = 0$ (from 6). 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 $[\underline{\mu}(\gamma_2), \overline{\mu}(\gamma_2)] \subset [\underline{\mu}(\gamma_1), \overline{\mu}(\gamma_1)]$.

Now we want to show,

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

First we calculate the limit of $t^D(\mu_0)$ when $\gamma \to 0$. Plugging $Q^{-1}(\mu_0) = \frac{\gamma}{\tilde{\mu}(\mu_0)}$ into equation (18) and multiplying by γ gives:

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

If $\gamma \to 0$ and μ_0 does not converge to 0 or 1, the above equation converges to $v_h - Q^{-1}(\mu_0) = v_l$.[34] 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 $\underline{\mu} = \frac{1}{2}\left(1 - \sqrt{1 - 4\gamma/(v_h - v_l)}\right)$,

[34] $\lim_{\gamma \to 0} \gamma \log \frac{\gamma}{Q^{-1}(\mu_0) - \gamma} = 0$

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 $\mu_0 v_h$. □

Proof of Proposition 4. Given $\lambda_P>\lambda_B$, we can rewrite the buyer's (IR) constraint,

$$V_P(\mu_0,s,x_r,t_r+t_u;\lambda_P)\ge V^0(\mu_0,s;\lambda_B), \tag{IR-P}$$

where $s\equiv v_h-t_b$. Similar as Lemma 1, we want to show that (IR-P) binds. Suppose (IR-P) holds strictly. Rewrite the buyer's Bellman equation for purchase,

$$\begin{aligned}V_P(\mu(\tau),s,x_r,t_r+t_u;\lambda_P)=\max\big\{&\mathbb{E}[v|\mu(\tau)]-(v_h-s),\mathbb{E}[v|\mu(\tau)]x_r-(t_r+t_u),\\&-kd\tau+\mu(\tau)\lambda_P d\tau s+(1-\mu(\tau)\lambda_P d\tau)V_P(\mu(\tau+d\tau),s,x_r,t_r+t_u;\lambda_P)\big\}.\end{aligned}\tag{25}$$

This leads to a differential equation conditional on learning,

$$U_P(\mu,s;\lambda_P)=s-\frac{k}{\mu\lambda_P}-(1-\mu)U_P'(\mu,s;\lambda_P), \tag{26}$$

where U_P' represents the partial derivative w.r.t to μ. Suppose the buyer stops learning at belief μ given the mechanism. Then by standard smooth pasting and value matching condition, the return revenue satisfies,

$$\begin{aligned}t_r+t_u&=\mathbb{E}[v|\mu]\frac{V_P'(\mu,s,x_r,t_r+t_u;\lambda_P)}{v_h-v_l}-V_P(\mu,s,x_r,t_r+t_u;\lambda_P)\\&=\frac{v_h}{v_h-v_l}V_P'-s+\frac{k}{\mu\lambda_P}.\end{aligned}\tag{27}$$

The second equality comes from the fact that V_P also satisfies the differential equation (26). Thus, a lower V_P implies a higher V_P', which implies a higher return revenue. Hence, if (IR-P) holds strictly, the seller can gain larger expected revenue by decreasing V_P while letting the buyer preserve the same stopping belief.

Since (IR-P) binds, under the optimal refund mechanism, $V_P(\mu_0, s, x_r, t_r + t_u; \lambda_P) = V^0(\mu_0, s; \lambda_B)$. Thus, imposing the boundary point $(\mu_0, V^0(\mu_0, s; \lambda_B))$ to the solution of U_P, we pin down the buyer's value function V_P conditional on learning. Let $t_u \equiv -V_P(\hat{\mu})$ where $\hat{\mu}$ is the belief at which $V_P' = 0$. Then normalizing the Bellman equation (25) by adding t_u at both sides, we obtain,

$$V_P(\mu(\tau), s, x_r, t_r + t_u; \lambda_P) + t_u = \max\big\{\mathbb{E}[v|\mu(\tau)] - (v_h - (s + t_u)), \mathbb{E}[v|\mu(\tau)]x_r - t_r,$$
$$-kd\tau + \mu(\tau)\lambda_P d\tau(s + t_u) + (1 - \mu(\tau)\lambda_P d\tau)(V_P(\mu(\tau + d\tau), s, x_r, t_r + t_u; \lambda_P) + t_u)\big\}.$$

By normalization,

$$V_P(\mu, s, x_r, t_r + t_u; \lambda_P) + t_u = V(\mu, s + t_u; \lambda_P), \tag{28}$$

which implies,

$$V(\mu_0, s + t_u; \lambda_P) - t_u = V_P(\mu_0, s, x_r, t_r + t_u; \lambda_P) = V^0(\mu_0, s; \lambda_B).$$

□

Proof of Proposition 5. Let $q_B^{-1}(\mu_0) := \frac{k}{\lambda_B \mu_0}$ be the inverse of quitting belief for before-transaction learning. Similarly, let $q_P^{-1}(\mu_0) := \frac{k}{\lambda_P \mu_0}$ and $Q_B^{-1}(\mu_0) = Q^{-1}(\mu_0; \lambda_B)$. There are two cases. (1) If $s \in [q_P^{-1}(\mu_0), q_B^{-1}(\mu_0))$, $V(\mu_0, s + t_u; \lambda_P) - t_u = V^0(\mu_0, s; \lambda_B) = 0$. (2) If $s > q_B^{-1}(\mu_0)$, $V(\mu_0, s + t_u; \lambda_P) - t_u = V^0(\mu_0, s; \lambda_B) > 0$. We discuss them separately.

Case (1). $s \in [q_P^{-1}(\mu_0), q_B^{-1}(\mu_0))$. Substituting $V(\mu_0, s + t_u; \lambda_P) - t_u = 0$ and equation (28) into equation (27), we obtain:[35]

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

Obviously,

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

[35]With our construction of t_u, t_r only depends on the stopping belief and s.

and the seller's expected revenue,

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

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

Case (2). $s \in [q_B^{-1}(\mu_0)), Q_B^{-1}(\mu_0))$. With a similar approach, we obtain:

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

Observe that

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

Then 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 $s \le \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$, otherwise, it is decreasing in s if $s > \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$.

Case (2a). If $\mu_0 \le v_l/v_h$, 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. Case (2b). 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 $q_B^{-1}(\mu_0) < \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)} \le Q_B^{-1}(\mu_0)$, rendering the optimal solution $s = \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$, which leads to the Stochastic Return in Proposition 5. Otherwise, if $\mu_0 \notin [\mu_0', \mu_0'']$, the optimal solution is $s = Q_B^{-1}(\mu_0)$.

Summarizing the above cases, if $q_B^{-1}(\mu_0)$ is the globally optimal s, then $\mu_0 \le v_l/v_h$. Notice that at $\mu_0 = \mu_0'$, the deterring learning revenue with $s = Q_B^{-1}(\mu_0)$ is higher than the optimal Stochastic Return with $s = \frac{k(v_h - v_l)}{\lambda_B(\mu_0 v_h - v_l)}$, which is higher than the revenue for $s = q_B^{-1}(\mu_0)$ given that the return policy is optimally chosen. Notice that the seller's expected revenue is continuous in μ_0 conditional on each case. Moreover, since the revenue for $s = q_B^{-1}(\mu_0)$ crosses the revenue for deterring learning at most three times on $\mu_0 \in [\underline{\mu}_0, \overline{\mu}_0]$ (cross from above at the first time at $\underline{\mu}_0$ and from below the last time at $\overline{\mu}_0$). Therefore if $\mu_0 < \mu_0'$, then setting s equal $q_B^{-1}(\mu_0)$ is sub-optimal. □

Proof of Proposition 6. If $4k < (v_h - v_l)\rho$, let $\underline{\mu}^N$ and $\overline{\mu}^N$ be the two roots such that the inequality,

$$\mathbb{E}[v|\mu]-G^{-1}(\mu)\geq 0,$$

is binding. That is, if the buyer stops learning at belief μ, consuming the item is weakly better than than walking away. This determines the interval of beliefs such that learning might be valuable.

Lemma 4. $\Pi^N(\mu)$ *is quasi-convex in* μ *for* $\mu_0\in[\underline{\mu}^N,\overline{\mu}^N]$.

Proof. With standard analysis, we obtain the expression of buyer's value function given an upper stopping belief μ,

$$V^N(\mu_0,G^{-1}(\mu))=-\frac{k(1-\mu+\mu_0)}{\rho(1-\mu)}+\mu_0(v_h-v_l)+\frac{k\mu_0}{\rho}\left(\log\left[\frac{1-\mu}{\mu}\right]-\log\left[\frac{1-\mu_0}{\mu_0}\right]\right).$$

With which, we can back out the return transfer t_r by the first constraint of program (N). Substituting it back to the objective function, we obtain,

$$\Pi^N(\mu):=\frac{\mu_0}{\mu}G^{-1}(\mu)+\frac{\mu-\mu_0}{\mu}\frac{v_l}{\mu_0(v_h-v_l)}V^N(\mu_0,\mu).$$

Take first order derivative of Π^N w.r.t μ,

$$\frac{d\Pi^N(\mu)}{d\mu}=R\underbrace{\left\{(2\mu-1)\mu_0 l+(\mu_0-1-\mu(\mu+\mu_0-1))+(1-\mu)^2\mu_0\left(\log\left[\frac{1-\mu}{\mu}\right]-\log\left[\frac{1-\mu_0}{\mu_0}\right]\right)\right\}}_{H(\mu)},$$

where $R=\frac{kv_l}{\rho(1-\mu)^2\mu^2(v_h-v_l)}>0$ and $l=\frac{v_h}{v_l}$. Thus, if $H(\mu)>0$, then $\Pi^N(\mu)$ is increasing in μ. Otherwise, $\Pi^N(\mu)$ is decreasing in μ.

Case 1, $\mu<1/2$. We can verify that $H(\mu)<0$ under this case. That is, if the upper stopping belief is smaller than 1/2, then $\Pi^N(\mu)$ is decreasing in μ, rendering μ_0 the optimal stopping belief.

Case 2, $\mu>1/2$. Notice that,

$$\frac{dH}{d\mu}=1-2\mu+2\mu_0 l-\frac{\mu_0}{\mu}+2(\mu-1)\mu_0\left(\log\left[\frac{1-\mu}{\mu}\right]-\log\left[\frac{1-\mu_0}{\mu_0}\right]\right).$$

Denote $J(\mu):=\frac{\mu-2\mu^2-\mu_0+2\mu\mu_0 l}{2\mu\mu_0-2\mu^2\mu_0}$ as the solution to $1-2\mu+2\mu_0 l-\frac{\mu_0}{\mu}+2(\mu-1)\mu_0 J(\mu)=0$. We can verify that $J(\mu)>\log\left[\frac{1-\mu}{\mu}\right]-\log\left[\frac{1-\mu_0}{\mu_0}\right]$ if $\mu=\frac{1}{2}$, while $J(\mu)<\log\left[\frac{1-\mu}{\mu}\right]-\log\left[\frac{1-\mu_0}{\mu_0}\right]$ if $\mu\to 0$. Note that, $J'(\mu)=\frac{\mu_0-2\mu\mu_0+\mu^2(2\mu_0 l-1)}{2(1-\mu)^2\mu^2\mu_0}$. Let $J'=0$, we obtain two roots, $\mu_1=\frac{\mu_0-\sqrt{\mu_0(1+\mu_0-2\mu_0 l)}}{2\mu_0 l-1}$ and $\mu_2=\frac{\mu_0+\sqrt{\mu_0(1+\mu_0-2\mu_0 l)}}{2\mu_0 l-1}$.

Case (2a), suppose $1+\mu_0-2\mu_0 l<0$. Then there is no real solutions. We can verify that under this case, $J'>0$. Thus, J is increasing and it can only cross $\log\left[\frac{1-\mu}{\mu}\right]-\log\left[\frac{1-\mu_0}{\mu_0}\right]$ once from below (the crossing belief is smaller than 1/2). It implies that $J(\mu)>\log\left[\frac{1-\mu}{\mu}\right]-\log\left[\frac{1-\mu_0}{\mu_0}\right]$ when $\mu\geq 1/2$, which further implies that $H'>0$. Thus, $H(\mu)$ can only cross 0 once from below. This means that Π^N is quasi-convex in μ for $\mu\geq 1/2$. Case (2b), suppose $1+\mu_0-2\mu_0 l\geq 0$. Then two roots exist. However, under this case, $H(\mu)<0$, therefore Π^N is decreasing in μ. □

The above Lemma implies that either Free Return that inducing upper stopping belief $\overline{G}(\mu_0)$ or Learning Deterrence that inducing μ_0 is optimal. Now we prove Proposition 6. Note that the revenue from Free Return is $\frac{\mu_0}{\overline{G}(\mu_0)}G^{-1}(\overline{G}(\mu_0))$ while Learning Deterrence is $G^{-1}(\mu_0)$.

Let $\Lambda(\mu_0)=\{\mu\neq\mu_0 : G^{-1}(\mu_0)=\frac{\mu_0}{\mu}G^{-1}(\mu)\}$. Thus,

$$\Lambda(\mu_0)=\frac{(1-\mu_0)(k+\rho v_l)}{k+\rho(1-\mu_0)v_l}.$$

We can verify that $\Lambda(\mu_0)$ is decreasing in μ_0. Recall that $\overline{G}(\mu_0)$ is the solution such that $V^N(\mu_0, G^{-1}(\mu))=0$. Take total differentiation we can verify that

$$\overline{G}'(\mu_0)=\frac{(1-\overline{G}(\mu_0))^2\overline{G}(\mu_0)}{\mu_0^2(1-\mu_0)}>0.$$

Therefore, $\Lambda(\mu_0)$ can cross $\overline{G}(\mu_0)$ at most once from above when $\mu_0\in[\underline{\mu}^N,\overline{\mu}^N]$.This implies that either there is an interval of beliefs such that the sender optimally chooses Free Return, or deterring learning is optimal for all $\mu_0\in[\underline{\mu}^N,\overline{\mu}^N]$. The existence of γ^{**} is driven by that $\overline{G}(\overline{\mu}^N)-\Lambda(\overline{\mu}^N)$ is decreasing in $\frac{k}{\rho}$. Moreover, if $\frac{k}{\rho}\to 0$ and $\overline{\mu}^N\to 1$, $\Lambda(1)=0<\overline{G}(1)$ and if $\frac{k}{\rho}\to\frac{v_h-v_l}{4}$ and $\overline{\mu}^N\to\frac{1}{2}$, $\Lambda(\frac{1}{2})=\frac{v_h+3v_l}{2(v_h+v_l)}>\overline{G}(\frac{1}{2})=\frac{1}{2}$.

□

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.

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.

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.

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.

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

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

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